

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER T

(19) World Intellectual Property Organization  
International Bureau



EXPRESS MAIL NO.  
EV020661113US

(43) International Publication Date  
2 August 2001 (02.08.2001)

PCT

(10) International Publication Number  
WO 01/54477 A2

(51) International Patent Classification: Not classified

(21) International Application Number: PCT/US01/02687

(22) International Filing Date: 25 January 2001 (25.01.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

09/491,404	25 January 2000 (25.01.2000)	US
09/617,746	17 July 2000 (17.07.2000)	US
09/631,451	3 August 2000 (03.08.2000)	US
09/663,870	15 September 2000 (15.09.2000)	US

(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier applications:

US	09/491,404 (CIP)
Filed on	25 January 2000 (25.01.2000)
US	09/617,746 (CIP)
Filed on	17 July 2000 (17.07.2000)
US	09/631,451 (CIP)
Filed on	3 August 2000 (03.08.2000)
US	09/663,870 (CIP)
Filed on	15 September 2000 (15.09.2000)

(71) Applicant (for all designated States except US): HYSEQ, INC. [US/US]; 670 Almanor Avenue, Sunnyvale, CA 94086 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): TANG, Y., Tom [US/US]; 4230 Ranwick Court, San Jose, CA 95118 (US). LIU, Chenghua [CN/US]; 1125 Ranchero Way #14, San Jose, CA 95117 (US). ZHOU, Ping [CN/US]; 1461 Japaur Lane, San Jose, CA 95132 (US). QIAN, Xiaohong, B. [CN/US]; 3662 Tumble Way, San Jose, CA 95132 (US). WANG, Zhiwei [CN/US]; 836 Alturas Avenue #B36,

Sunnyvale, CA 94085 (US). CHEN, Rui-Hong [US/US]; 1031 Flying Fish Street, Foster City, CA 94404 (US). ASUNDI, Vinod [US/US]; 709 Foster City Boulevard, Foster City, CA 94404 (US). CAO, Yicheng [CN/US]; 260 North Mathilda Avenue, Sunnyvale, CA 95086 (US). DRMANAC, Radoje, A. [YU/US]; 850 East Greenwich Place, Palo Alto, CA 94303 (US). ZHANG, Jie [CN/US]; 20800 Homestead Road #38B, Cupertino, CA 95014 (US). WERHMAN, Tom [US/US]; 300 Pasteur Drive, Edwards, R314, Stanford University Medical Center, Stanford, CA 94035 (US).

(74) Agent: ELRIFI, Ivor, R.; Mintz, Levin, Cohn, Ferris, Glovsky and Popeo, P.C., One Financial Center, Boston, MA 02111 (US).

(81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CI, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

— with declaration under Article 17(2)(a); without classification and without abstract; title not checked by the International Searching Authority

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.



WO 01/54477 A2

(54) Title: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES

(57) Abstract:

# NOVEL NUCLEIC ACIDS AND POLYPEPTIDES

## 1. TECHNICAL FIELD

The present invention provides novel polynucleotides and proteins encoded by such 5 polynucleotides, along with uses for these polynucleotides and proteins, for example in therapeutic, diagnostic and research methods.

## 2. BACKGROUND

Technology aimed at the discovery of protein factors (including *e.g.*, cytokines, such as 10 lymphokines, interferons, CSFs, chemokines, and interleukins) has matured rapidly over the past decade. The now routine hybridization cloning and expression cloning techniques clone novel polynucleotides "directly" in the sense that they rely on information directly related to the discovered protein (*i.e.*, partial DNA/amino acid sequence of the protein in the case of hybridization cloning; activity of the protein in the case of expression cloning). More recent 15 "indirect" cloning techniques such as signal sequence cloning, which isolates DNA sequences based on the presence of a now well-recognized secretory leader sequence motif, as well as various PCR-based or low stringency hybridization-based cloning techniques, have advanced the state of the art by making available large numbers of DNA/amino acid sequences for proteins that are known to have biological activity, for example, by virtue of their secreted nature in the 20 case of leader sequence cloning, by virtue of their cell or tissue source in the case of PCR-based techniques, or by virtue of structural similarity to other genes of known biological activity.

Identified polynucleotide and polypeptide sequences have numerous applications in, for example, diagnostics, forensics, gene mapping; identification of mutations responsible for 25 genetic disorders or other traits, to assess biodiversity, and to produce many other types of data and products dependent on DNA and amino acid sequences.

## 3. SUMMARY OF THE INVENTION

The compositions of the present invention include novel isolated polypeptides, novel 30 isolated polynucleotides encoding such polypeptides, including recombinant DNA molecules, cloned genes or degenerate variants thereof, especially naturally occurring variants such as allelic variants, antisense polynucleotide molecules, and antibodies that specifically recognize one or more epitopes present on such polypeptides, as well as hybridomas producing such antibodies.

The compositions of the present invention additionally include vectors, including expression 35 vectors, containing the polynucleotides of the invention, cells genetically engineered to contain such polynucleotides and cells genetically engineered to express such polynucleotides.

The present invention relates to a collection or library of at least one novel nucleic acid sequence assembled from expressed sequence tags (ESTs) isolated mainly by sequencing by hybridization (SBH), and in some cases, sequences obtained from one or more public databases. The invention relates also to the proteins encoded by such polynucleotides, along with therapeutic, 5 diagnostic and research utilities for these polynucleotides and proteins. These nucleic acid sequences are designated as SEQ ID NO: 1-1009. The polypeptides sequences are designated SEQ ID NO: 1010-2018. The nucleic acids and polypeptides are provided in the Sequence Listing. In the nucleic acids provided in the Sequence Listing, A is adenosine; C is cytosine; G is guanine; T is thymine; and N is any of the four bases. In the amino acids provided in the Sequence Listing, \* 10 corresponds to the stop codon.

The nucleic acid sequences of the present invention also include, nucleic acid sequences that hybridize to the complement of SEQ ID NO:1-1009 under stringent hybridization conditions; nucleic acid sequences which are allelic variants or species homologues of any of the nucleic acid sequences recited above, or nucleic acid sequences that encode a peptide comprising a specific 15 domain or truncation of the peptides encoded by SEQ ID NO:1-1009. A polynucleotide comprising a nucleotide sequence having at least 90% identity to an identifying sequence of SEQ ID NO:1-1009 or a degenerate variant or fragment thereof. The identifying sequence can be 100 base pairs in length.

The nucleic acid sequences of the present invention also include the sequence information 20 from the nucleic acid sequences of SEQ ID NO:1-1009. The sequence information can be a segment of any one of SEQ ID NO:1-1009 that uniquely identifies or represents the sequence information of SEQ ID NO:1-1009.

A collection as used in this application can be a collection of only one polynucleotide. The collection of sequence information or identifying information of each sequence can be provided on 25 a nucleic acid array. In one embodiment, segments of sequence information is provided on a nucleic acid array to detect the polynucleotide that contains the segment. The array can be designed to detect full-match or mismatch to the polynucleotide that contains the segment. The collection can also be provided in a computer-readable format.

This invention also includes the reverse or direct complement of any of the nucleic acid 30 sequences recited above; cloning or expression vectors containing the nucleic acid sequences; and host cells or organisms transformed with these expression vectors. Nucleic acid sequences (or their reverse or direct complements) according to the invention have numerous applications in a variety of techniques known to those skilled in the art of molecular biology, such as use as hybridization probes, use as primers for PCR, use in an array, use in computer-readable media, use in sequencing

full-length genes, use for chromosome and gene mapping, use in the recombinant production of protein, and use in the generation of anti-sense DNA or RNA, their chemical analogs and the like.

In a preferred embodiment, the nucleic acid sequences of SEQ ID NO:1-1009 or novel segments or parts of the nucleic acids of the invention are used as primers in expression assays that 5 are well known in the art. In a particularly preferred embodiment, the nucleic acid sequences of SEQ ID NO:1-1009 or novel segments or parts of the nucleic acids provided herein are used in diagnostics for identifying expressed genes or, as well known in the art and exemplified by Vollrath et al., *Science* 258:52-59 (1992), as expressed sequence tags for physical mapping of the human genome.

10 The isolated polynucleotides of the invention include, but are not limited to, a polynucleotide comprising any one of the nucleotide sequences set forth in SEQ ID NO:1-1009; a polynucleotide comprising any of the full length protein coding sequences of SEQ ID NO:1 - 1009; and a polynucleotide comprising any of the nucleotide sequences of the mature protein coding sequences of SEQ ID NO: 1- 1009. The polynucleotides of the present invention also include, but 15 are not limited to, a polynucleotide that hybridizes under stringent hybridization conditions to (a) the complement of any one of the nucleotide sequences set forth in SEQ ID NO:1-1009; (b) a nucleotide sequence encoding any one of the amino acid sequences set forth in the Sequence Listing (e.g., SEQ ID NO: 1010-2018); (c) a polynucleotide which is an allelic variant of any polynucleotides recited above; (d) a polynucleotide which encodes a species homolog (e.g. 20 orthologs) of any of the proteins recited above; or (e) a polynucleotide that encodes a polypeptide comprising a specific domain or truncation of any of the polypeptides comprising an amino acid sequence set forth in the Sequence Listing.

The isolated polypeptides of the invention include, but are not limited to, a polypeptide comprising any of the amino acid sequences set forth in the Sequence Listing; or the corresponding 25 full length or mature protein. Polypeptides of the invention also include polypeptides with biological activity that are encoded by (a) any of the polynucleotides having a nucleotide sequence set forth in SEQ ID NO:1-1009; or (b) polynucleotides that hybridize to the complement of the polynucleotides of (a) under stringent hybridization conditions. Biologically or immunologically active variants of any of the polypeptide sequences in the Sequence Listing, and "substantial 30 equivalents" thereof (e.g., with at least about 65%, 70%, 75%, 80%, 85%, 90%, 95%, 98% or 99% amino acid sequence identity) that preferably retain biological activity are also contemplated. The polypeptides of the invention may be wholly or partially chemically synthesized but are preferably produced by recombinant means using the genetically engineered cells (e.g. host cells) of the invention.

The invention also provides compositions comprising a polypeptide of the invention. Polypeptide compositions of the invention may further comprise an acceptable carrier, such as a hydrophilic, *e.g.*, pharmaceutically acceptable, carrier.

5 The invention also provides host cells transformed or transfected with a polynucleotide of the invention.

The invention also relates to methods for producing a polypeptide of the invention comprising growing a culture of the host cells of the invention in a suitable culture medium under conditions permitting expression of the desired polypeptide, and purifying the polypeptide from the culture or from the host cells. Preferred embodiments include those in which the 10 protein produced by such process is a mature form of the protein.

Polynucleotides according to the invention have numerous applications in a variety of techniques known to those skilled in the art of molecular biology. These techniques include use as hybridization probes, use as oligomers, or primers, for PCR, use for chromosome and gene mapping, use in the recombinant production of protein, and use in generation of anti-sense DNA 15 or RNA, their chemical analogs and the like. For example, when the expression of an mRNA is largely restricted to a particular cell or tissue type, polynucleotides of the invention can be used as hybridization probes to detect the presence of the particular cell or tissue mRNA in a sample using, *e.g.*, *in situ* hybridization.

20 In other exemplary embodiments, the polynucleotides are used in diagnostics as expressed sequence tags for identifying expressed genes or, as well known in the art and exemplified by Vollrath et al., *Science* 258:52-59 (1992), as expressed sequence tags for physical mapping of the human genome.

25 The polypeptides according to the invention can be used in a variety of conventional procedures and methods that are currently applied to other proteins. For example, a polypeptide of the invention can be used to generate an antibody that specifically binds the polypeptide. Such antibodies, particularly monoclonal antibodies, are useful for detecting or quantitating the polypeptide in tissue. The polypeptides of the invention can also be used as molecular weight markers, and as a food supplement.

30 Methods are also provided for preventing, treating, or ameliorating a medical condition which comprises the step of administering to a mammalian subject a therapeutically effective amount of a composition comprising a polypeptide of the present invention and a pharmaceutically acceptable carrier.

35 In particular, the polypeptides and polynucleotides of the invention can be utilized, for example, in methods for the prevention and/or treatment of disorders involving aberrant protein expression or biological activity.

The present invention further relates to methods for detecting the presence of the polynucleotides or polypeptides of the invention in a sample. Such methods can, for example, be utilized as part of prognostic and diagnostic evaluation of disorders as recited herein and for the identification of subjects exhibiting a predisposition to such conditions. The invention provides

5 a method for detecting the polynucleotides of the invention in a sample, comprising contacting the sample with a compound that binds to and forms a complex with the polynucleotide of interest for a period sufficient to form the complex and under conditions sufficient to form a complex and detecting the complex such that if a complex is detected, the polynucleotide of interest is detected. The invention also provides a method for detecting the polypeptides of the

10 invention in a sample comprising contacting the sample with a compound that binds to and forms a complex with the polypeptide under conditions and for a period sufficient to form the complex and detecting the formation of the complex such that if a complex is formed, the polypeptide is detected.

The invention also provides kits comprising polynucleotide probes and/or monoclonal antibodies, and optionally quantitative standards, for carrying out methods of the invention. Furthermore, the invention provides methods for evaluating the efficacy of drugs, and monitoring the progress of patients, involved in clinical trials for the treatment of disorders as recited above.

The invention also provides methods for the identification of compounds that modulate (i.e., increase or decrease) the expression or activity of the polynucleotides and/or polypeptides of the invention. Such methods can be utilized, for example, for the identification of compounds that can ameliorate symptoms of disorders as recited herein. Such methods can include, but are not limited to, assays for identifying compounds and other substances that interact with (e.g., bind to) the polypeptides of the invention. The invention provides a method for identifying a compound that binds to the polypeptides of the invention comprising contacting the compound with a polypeptide of the invention in a cell for a time sufficient to form a polypeptide/compound complex, wherein the complex drives expression of a reporter gene sequence in the cell; and detecting the complex by detecting the reporter gene sequence expression such that if expression of the reporter gene is detected the compound binds to a polypeptide of the invention is identified.

The methods of the invention also provides methods for treatment which involve the administration of the polynucleotides or polypeptides of the invention to individuals exhibiting symptoms or tendencies. In addition, the invention encompasses methods for treating diseases or disorders as recited herein comprising administering compounds and other substances that modulate the overall activity of the target gene products. Compounds and other substances can

effect such modulation either on the level of target gene/protein expression or target protein activity.

The polypeptides of the present invention and the polynucleotides encoding them are also useful for the same functions known to one of skill in the art as the polypeptides and 5 polynucleotides to which they have homology (set forth in Table 2). If no homology is set forth for a sequence, then the polypeptides and polynucleotides of the present invention are useful for a variety of applications, as described herein, including use in arrays for detection.

10

#### 4. DETAILED DESCRIPTION OF THE INVENTION

##### 4.1 DEFINITIONS

It must be noted that as used herein and in the appended claims, the singular forms "a", "an" and "the" include plural references unless the context clearly dictates otherwise.

15

The term "active" refers to those forms of the polypeptide which retain the biologic and/or immunologic activities of any naturally occurring polypeptide. According to the invention, the terms "biologically active" or "biological activity" refer to a protein or peptide having structural, regulatory or biochemical functions of a naturally occurring molecule. Likewise "immunologically active" or "immunological activity" refers to the capability of the 20 natural, recombinant or synthetic polypeptide to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies.

The term "activated cells" as used in this application are those cells which are engaged in extracellular or intracellular membrane trafficking, including the export of secretory or enzymatic molecules as part of a normal or disease process.

25

The terms "complementary" or "complementarity" refer to the natural binding of polynucleotides by base pairing. For example, the sequence 5'-AGT-3' binds to the complementary sequence 3'-TCA-5'. Complementarity between two single-stranded molecules may be "partial" such that only some of the nucleic acids bind or it may be "complete" such that total complementarity exists between the single stranded molecules. The degree of 30 complementarity between the nucleic acid strands has significant effects on the efficiency and strength of the hybridization between the nucleic acid strands.

The term "embryonic stem cells (ES)" refers to a cell that can give rise to many differentiated cell types in an embryo or an adult, including the germ cells. The term "germ line stem cells (GSCs)" refers to stem cells derived from primordial stem cells that provide a steady 35 and continuous source of germ cells for the production of gametes. The term "primordial germ

cells (PGCs)" refers to a small population of cells set aside from other cell lineages particularly from the yolk sac, mesenteries, or gonadal ridges during embryogenesis that have the potential to differentiate into germ cells and other cells. PGCs are the source from which GSCs and ES cells are derived. The PGCs, the GSCs and the ES cells are capable of self-renewal. Thus these cells 5 not only populate the germ line and give rise to a plurality of terminally differentiated cells that comprise the adult specialized organs, but are able to regenerate themselves.

The term "expression modulating fragment," EMF, means a series of nucleotides which modulates the expression of an operably linked ORF or another EMF.

As used herein, a sequence is said to "modulate the expression of an operably linked 10 sequence" when the expression of the sequence is altered by the presence of the EMF. EMFs include, but are not limited to, promoters, and promoter modulating sequences (inducible elements). One class of EMFs are nucleic acid fragments which induce the expression of an operably linked ORF in response to a specific regulatory factor or physiological event.

The terms "nucleotide sequence" or "nucleic acid" or "polynucleotide" or 15 "oligonucleotide" are used interchangeably and refer to a heteropolymer of nucleotides or the sequence of these nucleotides. These phrases also refer to DNA or RNA of genomic or synthetic origin which may be single-stranded or double-stranded and may represent the sense or the antisense strand, to peptide nucleic acid (PNA) or to any DNA-like or RNA-like material. In the sequences herein A is adenine, C is cytosine, T is thymine, G is guanine and N is A, C, G or T 20 (U). It is contemplated that where the polynucleotide is RNA, the T (thymine) in the sequences provided herein is substituted with U (uracil). Generally, nucleic acid segments provided by this invention may be assembled from fragments of the genome and short oligonucleotide linkers, or from a series of oligonucleotides, or from individual nucleotides, to provide a synthetic nucleic acid which is capable of being expressed in a recombinant transcriptional unit comprising 25 regulatory elements derived from a microbial or viral operon, or a eukaryotic gene.

The terms "oligonucleotide fragment" or a "polynucleotide fragment", "portion," or "segment" or "probe" or "primer" are used interchangeably and refer to a sequence of nucleotide residues which are at least about 5 nucleotides, more preferably at least about 7 nucleotides, more preferably at least about 9 nucleotides, more preferably at least about 11 nucleotides and 30 most preferably at least about 17 nucleotides. The fragment is preferably less than about 500 nucleotides, preferably less than about 200 nucleotides, more preferably less than about 100 nucleotides, more preferably less than about 50 nucleotides and most preferably less than 30 nucleotides. Preferably the probe is from about 6 nucleotides to about 200 nucleotides, preferably from about 15 to about 50 nucleotides, more preferably from about 17 to 30 35 nucleotides and most preferably from about 20 to 25 nucleotides. Preferably the fragments can

be used in polymerase chain reaction (PCR), various hybridization procedures or microarray procedures to identify or amplify identical or related parts of mRNA or DNA molecules. A fragment or segment may uniquely identify each polynucleotide sequence of the present invention. Preferably the fragment comprises a sequence substantially similar to any one of SEQ 5 ID NOS:1-1009.

Probes may, for example, be used to determine whether specific mRNA molecules are present in a cell or tissue or to isolate similar nucleic acid sequences from chromosomal DNA as described by Walsh et al. (Walsh, P.S. et al., 1992, PCR Methods Appl 1:241-250). They may be labeled by nick translation, Klenow fill-in reaction, PCR, or other methods well known in the 10 art. Probes of the present invention, their preparation and/or labeling are elaborated in Sambrook, J. et al., 1989, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, NY; or Ausubel, F.M. et al., 1989, Current Protocols in Molecular Biology, John Wiley & Sons, New York NY, both of which are incorporated herein by reference in their entirety.

15 The nucleic acid sequences of the present invention also include the sequence information from the nucleic acid sequences of SEQ ID NO:1-1009. The sequence information can be a segment of any one of SEQ ID NO:1-1009 that uniquely identifies or represents the sequence information of that sequence of SEQ ID NO:1-1009. One such segment can be a twenty-mer nucleic acid sequence because the probability that a twenty-mer is fully matched in 20 the human genome is 1 in 300. In the human genome, there are three billion base pairs in one set of chromosomes. Because  $4^{20}$  possible twenty-mers exist, there are 300 times more twenty-mers than there are base pairs in a set of human chromosomes. Using the same analysis, the probability for a seventeen-mer to be fully matched in the human genome is approximately 1 in 5. When these segments are used in arrays for expression studies, fifteen-mer segments can be 25 used. The probability that the fifteen-mer is fully matched in the expressed sequences is also approximately one in five because expressed sequences comprise less than approximately 5% of the entire genome sequence.

Similarly, when using sequence information for detecting a single mismatch, a segment can be a twenty-five mer. The probability that the twenty-five mer would appear in a human genome 30 with a single mismatch is calculated by multiplying the probability for a full match ( $1 \div 4^{25}$ ) times the increased probability for mismatch at each nucleotide position ( $3 \times 25$ ). The probability that an eighteen mer with a single mismatch can be detected in an array for expression studies is approximately one in five. The probability that a twenty-mer with a single mismatch can be detected in a human genome is approximately one in five.

The term "open reading frame," ORF, means a series of nucleotide triplets coding for amino acids without any termination codons and is a sequence translatable into protein.

The terms "operably linked" or "operably associated" refer to functionally related nucleic acid sequences. For example, a promoter is operably associated or operably linked with a coding sequence if the promoter controls the transcription of the coding sequence. While operably linked nucleic acid sequences can be contiguous and in the same reading frame, certain genetic elements *e.g.* repressor genes are not contiguously linked to the coding sequence but still control transcription/translation of the coding sequence.

The term "pluripotent" refers to the capability of a cell to differentiate into a number of differentiated cell types that are present in an adult organism. A pluripotent cell is restricted in its differentiation capability in comparison to a totipotent cell.

The terms "polypeptide" or "peptide" or "amino acid sequence" refer to an oligopeptide, peptide, polypeptide or protein sequence or fragment thereof and to naturally occurring or synthetic molecules. A polypeptide "fragment," "portion," or "segment" is a stretch of amino acid residues of at least about 5 amino acids, preferably at least about 7 amino acids, more preferably at least about 9 amino acids and most preferably at least about 17 or more amino acids. The peptide preferably is not greater than about 200 amino acids, more preferably less than 150 amino acids and most preferably less than 100 amino acids. Preferably the peptide is from about 5 to about 200 amino acids. To be active, any polypeptide must have sufficient length to display biological and/or immunological activity.

The term "naturally occurring polypeptide" refers to polypeptides produced by cells that have not been genetically engineered and specifically contemplates various polypeptides arising from post-translational modifications of the polypeptide including, but not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation and acylation.

The term "translated protein coding portion" means a sequence which encodes for the full length protein which may include any leader sequence or any processing sequence.

The term "mature protein coding sequence" means a sequence which encodes a peptide or protein without a signal or leader sequence. The "mature protein portion" means that portion of the protein which does not include a signal or leader sequence. The peptide may have been produced by processing in the cell which removes any leader/signal sequence. The mature protein portion may or may not include the initial methionine residue. The methionine residue may be removed from the protein during processing in the cell. The peptide may be produced synthetically or the protein may have been produced using a polynucleotide only encoding for the mature protein coding sequence.

The term "derivative" refers to polypeptides chemically modified by such techniques as ubiquitination, labeling (*e.g.*, with radionuclides or various enzymes), covalent polymer attachment such as pegylation (derivatization with polyethylene glycol) and insertion or substitution by chemical synthesis of amino acids such as ornithine, which do not normally occur in human proteins.

The term "variant"(*or "analog"*) refers to any polypeptide differing from naturally occurring polypeptides by amino acid insertions, deletions, and substitutions, created using, *e.g.*, recombinant DNA techniques. Guidance in determining which amino acid residues may be replaced, added or deleted without abolishing activities of interest, may be found by comparing the sequence of the particular polypeptide with that of homologous peptides and minimizing the number of amino acid sequence changes made in regions of high homology (conserved regions) or by replacing amino acids with consensus sequence.

Alternatively, recombinant variants encoding these same or similar polypeptides may be synthesized or selected by making use of the "redundancy" in the genetic code. Various codon substitutions, such as the silent changes which produce various restriction sites, may be introduced to optimize cloning into a plasmid or viral vector or expression in a particular prokaryotic or eukaryotic system. Mutations in the polynucleotide sequence may be reflected in the polypeptide or domains of other peptides added to the polypeptide to modify the properties of any part of the polypeptide, to change characteristics such as ligand-binding affinities, interchain affinities, or degradation/turnover rate.

Preferably, amino acid "substitutions" are the result of replacing one amino acid with another amino acid having similar structural and/or chemical properties, *i.e.*, conservative amino acid replacements. "Conservative" amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues involved. For example, nonpolar (hydrophobic) amino acids include alanine, leucine, isoleucine, valine, proline, phenylalanine, tryptophan, and methionine; polar neutral amino acids include glycine, serine, threonine, cysteine, tyrosine, asparagine, and glutamine; positively charged (basic) amino acids include arginine, lysine, and histidine; and negatively charged (acidic) amino acids include aspartic acid and glutamic acid. "Insertions" or "deletions" are preferably in the range of about 1 to 20 amino acids, more preferably 1 to 10 amino acids. The variation allowed may be experimentally determined by systematically making insertions, deletions, or substitutions of amino acids in a polypeptide molecule using recombinant DNA techniques and assaying the resulting recombinant variants for activity.

Alternatively, where alteration of function is desired, insertions, deletions or non-conservative alterations can be engineered to produce altered polypeptides. Such alterations

can, for example, alter one or more of the biological functions or biochemical characteristics of the polypeptides of the invention. For example, such alterations may change polypeptide characteristics such as ligand-binding affinities, interchain affinities, or degradation/turnover rate. Further, such alterations can be selected so as to generate polypeptides that are better suited 5 for expression, scale up and the like in the host cells chosen for expression. For example, cysteine residues can be deleted or substituted with another amino acid residue in order to eliminate disulfide bridges.

The terms "purified" or "substantially purified" as used herein denotes that the indicated nucleic acid or polypeptide is present in the substantial absence of other biological 10 macromolecules, *e.g.*, polynucleotides, proteins, and the like. In one embodiment, the polynucleotide or polypeptide is purified such that it constitutes at least 95% by weight, more preferably at least 99% by weight, of the indicated biological macromolecules present (but water, buffers, and other small molecules, especially molecules having a molecular weight of less than 1000 daltons, can be present).

15 The term "isolated" as used herein refers to a nucleic acid or polypeptide separated from at least one other component (*e.g.*, nucleic acid or polypeptide) present with the nucleic acid or polypeptide in its natural source. In one embodiment, the nucleic acid or polypeptide is found in the presence of (if anything) only a solvent, buffer, ion, or other component normally present in a solution of the same. The terms "isolated" and "purified" do not encompass nucleic acids or 20 polypeptides present in their natural source.

The term "recombinant," when used herein to refer to a polypeptide or protein, means 25 that a polypeptide or protein is derived from recombinant (*e.g.*, microbial, insect, or mammalian) expression systems. "Microbial" refers to recombinant polypeptides or proteins made in bacterial or fungal (*e.g.*, yeast) expression systems. As a product, "recombinant microbial" defines a polypeptide or protein essentially free of native endogenous substances and unaccompanied by associated native glycosylation. Polypeptides or proteins expressed in most bacterial cultures, *e.g.*, *E. coli*, will be free of glycosylation modifications; polypeptides or proteins expressed in yeast will have a glycosylation pattern in general different from those expressed in mammalian cells.

30 The term "recombinant expression vehicle or vector" refers to a plasmid or phage or virus or vector, for expressing a polypeptide from a DNA (RNA) sequence. An expression vehicle can comprise a transcriptional unit comprising an assembly of (1) a genetic element or elements having a regulatory role in gene expression, for example, promoters or enhancers, (2) a structural or coding sequence which is transcribed into mRNA and translated into protein, and (3) 35 appropriate transcription initiation and termination sequences. Structural units intended for use

in yeast or eukaryotic expression systems preferably include a leader sequence enabling extracellular secretion of translated protein by a host cell. Alternatively, where recombinant protein is expressed without a leader or transport sequence, it may include an amino terminal methionine residue. This residue may or may not be subsequently cleaved from the expressed 5 recombinant protein to provide a final product.

The term "recombinant expression system" means host cells which have stably integrated a recombinant transcriptional unit into chromosomal DNA or carry the recombinant transcriptional unit extrachromosomally. Recombinant expression systems as defined herein will express heterologous polypeptides or proteins upon induction of the regulatory elements linked 10 to the DNA segment or synthetic gene to be expressed. This term also means host cells which have stably integrated a recombinant genetic element or elements having a regulatory role in gene expression, for example, promoters or enhancers. Recombinant expression systems as defined herein will express polypeptides or proteins endogenous to the cell upon induction of the regulatory elements linked to the endogenous DNA segment or gene to be expressed. The cells 15 can be prokaryotic or eukaryotic.

The term "secreted" includes a protein that is transported across or through a membrane, including transport as a result of signal sequences in its amino acid sequence when it is expressed in a suitable host cell. "Secreted" proteins include without limitation proteins secreted wholly (e.g., soluble proteins) or partially (e.g., receptors) from the cell in which they are 20 expressed. "Secreted" proteins also include without limitation proteins that are transported across the membrane of the endoplasmic reticulum. "Secreted" proteins are also intended to include proteins containing non-typical signal sequences (e.g. Interleukin-1 Beta, see Krasney, P.A. and Young, P.R. (1992) Cytokine 4(2):134 -143) and factors released from damaged cells (e.g. Interleukin-1 Receptor Antagonist, see Arend, W.P. et. al. (1998) Annu. Rev. Immunol. 25 16:27-55)

Where desired, an expression vector may be designed to contain a "signal or leader sequence" which will direct the polypeptide through the membrane of a cell. Such a sequence may be naturally present on the polypeptides of the present invention or provided from heterologous protein sources by recombinant DNA techniques.

30 The term "stringent" is used to refer to conditions that are commonly understood in the art as stringent. Stringent conditions can include highly stringent conditions (*i.e.*, hybridization to filter-bound DNA in 0.5 M NaHPO<sub>4</sub>, 7% sodium dodecyl sulfate (SDS), 1 mM EDTA at 65°C, and washing in 0.1X SSC/0.1% SDS at 68°C), and moderately stringent conditions (*i.e.*, washing in 0.2X SSC/0.1% SDS at 42°C). Other exemplary hybridization conditions are 35 described herein in the examples.

In instances of hybridization of deoxyoligonucleotides, additional exemplary stringent hybridization conditions include washing in 6X SSC/0.05% sodium pyrophosphate at 37°C (for 14-base oligonucleotides), 48°C (for 17-base oligos), 55°C (for 20-base oligonucleotides), and 60°C (for 23-base oligonucleotides).

5 As used herein, "substantially equivalent" can refer both to nucleotide and amino acid sequences, for example a mutant sequence, that varies from a reference sequence by one or more substitutions, deletions, or additions, the net effect of which does not result in an adverse functional dissimilarity between the reference and subject sequences. Typically, such a substantially equivalent sequence varies from one of those listed herein by no more than about 10 35% (*i.e.*, the number of individual residue substitutions, additions, and/or deletions in a substantially equivalent sequence, as compared to the corresponding reference sequence, divided by the total number of residues in the substantially equivalent sequence is about 0.35 or less). Such a sequence is said to have 65% sequence identity to the listed sequence. In one embodiment, a substantially equivalent, *e.g.*, mutant, sequence of the invention varies from a 15 listed sequence by no more than 30% (70% sequence identity); in a variation of this embodiment, by no more than 25% (75% sequence identity); and in a further variation of this embodiment, by no more than 20% (80% sequence identity) and in a further variation of this embodiment, by no more than 10% (90% sequence identity) and in a further variation of this embodiment, by no more than 5% (95% sequence identity). Substantially equivalent, *e.g.*, 20 mutant, amino acid sequences according to the invention preferably have at least 80% sequence identity with a listed amino acid sequence, more preferably at least 85% sequence identity, more preferably at least 90% sequence identity, more preferably at least 95% identity, more preferably at least 98% identity, and most preferably at least 99% identity. Substantially equivalent nucleotide sequences of the invention can have lower percent sequence identities, taking into 25 account, for example, the redundancy or degeneracy of the genetic code. Preferably, nucleotide sequence has at least about 65% identity, more preferably at least about 75% identity, more preferably at least about 80% sequence identity, more preferably at least about 85% sequence identity, more preferably at least about 90% sequence identity, and most preferably at least about 95% identity, more preferably at least about 98% sequence identity, and most preferably at least 30 about 99% sequence identity. For the purposes of the present invention, sequences having substantially equivalent biological activity and substantially equivalent expression characteristics are considered substantially equivalent. For the purposes of determining equivalence, truncation of the mature sequence (*e.g.*, via a mutation which creates a spurious stop codon) should be disregarded. Sequence identity may be determined, *e.g.*, using the Jotun Hein method (Hein, J.

(1990) Methods Enzymol. 183:626-645). Identity between sequences can also be determined by other methods known in the art, *e.g.* by varying hybridization conditions.

The term "totipotent" refers to the capability of a cell to differentiate into all of the cell types of an adult organism.

5 The term "transformation" means introducing DNA into a suitable host cell so that the DNA is replicable, either as an extrachromosomal element, or by chromosomal integration. The term "transfection" refers to the taking up of an expression vector by a suitable host cell, whether or not any coding sequences are in fact expressed. The term "infection" refers to the introduction of nucleic acids into a suitable host cell by use of a virus or viral vector.

10 As used herein, an "uptake modulating fragment," UMF, means a series of nucleotides which mediate the uptake of a linked DNA fragment into a cell. UMFs can be readily identified using known UMFs as a target sequence or target motif with the computer-based systems described below. The presence and activity of a UMF can be confirmed by attaching the suspected UMF to a marker sequence. The resulting nucleic acid molecule is then incubated 15 with an appropriate host under appropriate conditions and the uptake of the marker sequence is determined. As described above, a UMF will increase the frequency of uptake of a linked marker sequence.

Each of the above terms is meant to encompass all that is described for each, unless the context dictates otherwise.

20

#### 4.2 NUCLEIC ACIDS OF THE INVENTION

Nucleotide sequences of the invention are set forth in the Sequence Listing.

The isolated polynucleotides of the invention include a polynucleotide comprising the nucleotide sequences of SEQ ID NO:1-1009 ; a polynucleotide encoding any one of the peptide 25 sequences of SEQ ID NO:1010-2018; and a polynucleotide comprising the nucleotide sequence encoding the mature protein coding sequence of the polypeptides of any one of SEQ ID NO:1010-2018. The polynucleotides of the present invention also include, but are not limited to, a polynucleotide that hybridizes under stringent conditions to (a) the complement of any of the nucleotides sequences of SEQ ID NO:1-1009 ; (b) nucleotide sequences encoding any one of the 30 amino acid sequences set forth in the Sequence Listing; (c) a polynucleotide which is an allelic variant of any polynucleotide recited above; (d) a polynucleotide which encodes a species homolog of any of the proteins recited above; or (e) a polynucleotide that encodes a polypeptide comprising a specific domain or truncation of the polypeptides of SEQ ID NO: 1010-2018. Domains of interest may depend on the nature of the encoded polypeptide; *e.g.*, domains in 35 receptor-like polypeptides include ligand-binding, extracellular, transmembrane, or cytoplasmic

domains, or combinations thereof; domains in immunoglobulin-like proteins include the variable immunoglobulin-like domains; domains in enzyme-like polypeptides include catalytic and substrate binding domains; and domains in ligand polypeptides include receptor-binding domains.

5 The polynucleotides of the invention include naturally occurring or wholly or partially synthetic DNA, *e.g.*, cDNA and genomic DNA, and RNA, *e.g.*, mRNA. The polynucleotides may include all of the coding region of the cDNA or may represent a portion of the coding region of the cDNA.

10 The present invention also provides genes corresponding to the cDNA sequences disclosed herein. The corresponding genes can be isolated in accordance with known methods using the sequence information disclosed herein. Such methods include the preparation of probes or primers from the disclosed sequence information for identification and/or amplification of genes in appropriate genomic libraries or other sources of genomic materials. Further 5' and 3' sequence can be obtained using methods known in the art. For example, full length cDNA or genomic DNA that 15 corresponds to any of the polynucleotides of SEQ ID NO:1-1009 can be obtained by screening appropriate cDNA or genomic DNA libraries under suitable hybridization conditions using any of the polynucleotides of SEQ ID NO:1-1009 or a portion thereof as a probe. Alternatively, the polynucleotides of SEQ ID NO:1-1009 may be used as the basis for suitable primer(s) that allow identification and/or amplification of genes in appropriate genomic DNA or cDNA libraries.

20 The nucleic acid sequences of the invention can be assembled from ESTs and sequences (including cDNA and genomic sequences) obtained from one or more public databases, such as dbEST, gbpri, and UniGene. The EST sequences can provide identifying sequence information, representative fragment or segment information, or novel segment information for the full-length gene.

25 The polynucleotides of the invention also provide polynucleotides including nucleotide sequences that are substantially equivalent to the polynucleotides recited above. Polynucleotides according to the invention can have, *e.g.*, at least about 65%, at least about 70%, at least about 75%, at least about 80%, 81%, 82%, 83%, 84%, more typically at least about 85%, 86%, 87%, 88%, 89%, more typically at least about 90%, 91%, 92%, 93%, 94%, and even more typically at 30 least about 95%, 96%, 97%, 98%, 99%, sequence identity to a polynucleotide recited above.

35 Included within the scope of the nucleic acid sequences of the invention are nucleic acid sequence fragments that hybridize under stringent conditions to any of the nucleotide sequences of SEQ ID NO:1-1009, or complements thereof, which fragment is greater than about 5 nucleotides, preferably 7 nucleotides, more preferably greater than 9 nucleotides and most preferably greater than 17 nucleotides. Fragments of, *e.g.* 15, 17, or 20 nucleotides or more that

are selective for (*i.e.* specifically hybridize to any one of the polynucleotides of the invention) are contemplated. Probes capable of specifically hybridizing to a polynucleotide can differentiate polynucleotide sequences of the invention from other polynucleotide sequences in the same family of genes or can differentiate human genes from genes of other species, and are 5 preferably based on unique nucleotide sequences.

The sequences falling within the scope of the present invention are not limited to these specific sequences, but also include allelic and species variations thereof. Allelic and species variations can be routinely determined by comparing the sequence provided SEQ ID NO:1-1009, a representative fragment thereof, or a nucleotide sequence at least 90% identical, preferably 95% 10 identical, to SEQ ID NO:1-1009 with a sequence from another isolate of the same species. Furthermore, to accommodate codon variability, the invention includes nucleic acid molecules coding for the same amino acid sequences as do the specific ORFs disclosed herein. In other words, in the coding region of an ORF, substitution of one codon for another codon that encodes the same amino acid is expressly contemplated.

15 The nearest neighbor or homology result for the nucleic acids of the present invention, including SEQ ID NO:1-1009, can be obtained by searching a database using an algorithm or a program. Preferably, a BLAST which stands for Basic Local Alignment Search Tool is used to search for local sequence alignments (Altshul, S.F. *J Mol. Evol.* 36 290-300 (1993) and Altschul S.F. et al. *J. Mol. Biol.* 21:403-410 (1990)). Alternatively a FASTA version 3 search against 20 Genpept, using Fastxy algorithm.

Species homologs (or orthologs) of the disclosed polynucleotides and proteins are also provided by the present invention. Species homologs may be isolated and identified by making suitable probes or primers from the sequences provided herein and screening a suitable nucleic acid source from the desired species.

25 The invention also encompasses allelic variants of the disclosed polynucleotides or proteins; that is, naturally-occurring alternative forms of the isolated polynucleotide which also encode proteins which are identical, homologous or related to that encoded by the polynucleotides.

The nucleic acid sequences of the invention are further directed to sequences which 30 encode variants of the described nucleic acids. These amino acid sequence variants may be prepared by methods known in the art by introducing appropriate nucleotide changes into a native or variant polynucleotide. There are two variables in the construction of amino acid sequence variants: the location of the mutation and the nature of the mutation. Nucleic acids encoding the amino acid sequence variants are preferably constructed by mutating the 35 polynucleotide to encode an amino acid sequence that does not occur in nature. These nucleic

acid alterations can be made at sites that differ in the nucleic acids from different species (variable positions) or in highly conserved regions (constant regions). Sites at such locations will typically be modified in series, *e.g.*, by substituting first with conservative choices (*e.g.*, hydrophobic amino acid to a different hydrophobic amino acid) and then with more distant choices (*e.g.*, hydrophobic amino acid to a charged amino acid), and then deletions or insertions may be made at the target site. Amino acid sequence deletions generally range from about 1 to 30 residues, preferably about 1 to 10 residues, and are typically contiguous. Amino acid insertions include amino- and/or carboxyl-terminal fusions ranging in length from one to one hundred or more residues, as well as intrasequence insertions of single or multiple amino acid residues. Intrasequence insertions may range generally from about 1 to 10 amino residues, preferably from 1 to 5 residues. Examples of terminal insertions include the heterologous signal sequences necessary for secretion or for intracellular targeting in different host cells and sequences such as FLAG or poly-histidine sequences useful for purifying the expressed protein.

In a preferred method, polynucleotides encoding the novel amino acid sequences are changed via site-directed mutagenesis. This method uses oligonucleotide sequences to alter a polynucleotide to encode the desired amino acid variant, as well as sufficient adjacent nucleotides on both sides of the changed amino acid to form a stable duplex on either side of the site of being changed. In general, the techniques of site-directed mutagenesis are well known to those of skill in the art and this technique is exemplified by publications such as, Edelman et al., *DNA* 2:183 (1983). A versatile and efficient method for producing site-specific changes in a polynucleotide sequence was published by Zoller and Smith, *Nucleic Acids Res.* 10:6487-6500 (1982). PCR may also be used to create amino acid sequence variants of the novel nucleic acids. When small amounts of template DNA are used as starting material, primer(s) that differs slightly in sequence from the corresponding region in the template DNA can generate the desired amino acid variant. PCR amplification results in a population of product DNA fragments that differ from the polynucleotide template encoding the polypeptide at the position specified by the primer. The product DNA fragments replace the corresponding region in the plasmid and this gives a polynucleotide encoding the desired amino acid variant.

A further technique for generating amino acid variants is the cassette mutagenesis technique described in Wells et al., *Gene* 34:315 (1985); and other mutagenesis techniques well known in the art, such as, for example, the techniques in Sambrook et al., *supra*, and *Current Protocols in Molecular Biology*, Ausubel et al. Due to the inherent degeneracy of the genetic code, other DNA sequences which encode substantially the same or a functionally equivalent amino acid sequence may be used in the practice of the invention for the cloning and expression

of these novel nucleic acids. Such DNA sequences include those which are capable of hybridizing to the appropriate novel nucleic acid sequence under stringent conditions.

Polynucleotides encoding preferred polypeptide truncations of the invention can be used to generate polynucleotides encoding chimeric or fusion proteins comprising one or more domains of the invention and heterologous protein sequences.

The polynucleotides of the invention additionally include the complement of any of the polynucleotides recited above. The polynucleotide can be DNA (genomic, cDNA, amplified, or synthetic) or RNA. Methods and algorithms for obtaining such polynucleotides are well known to those of skill in the art and can include, for example, methods for determining hybridization conditions that can routinely isolate polynucleotides of the desired sequence identities.

In accordance with the invention, polynucleotide sequences comprising the mature protein coding sequences corresponding to any one of SEQ ID NO:1-1009, or functional equivalents thereof, may be used to generate recombinant DNA molecules that direct the expression of that nucleic acid, or a functional equivalent thereof, in appropriate host cells. Also included are the cDNA inserts of any of the clones identified herein.

A polynucleotide according to the invention can be joined to any of a variety of other nucleotide sequences by well-established recombinant DNA techniques (see Sambrook J et al. (1989) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, NY). Useful nucleotide sequences for joining to polynucleotides include an assortment of vectors, *e.g.*, plasmids, cosmids, lambda phage derivatives, phagemids, and the like, that are well known in the art. Accordingly, the invention also provides a vector including a polynucleotide of the invention and a host cell containing the polynucleotide. In general, the vector contains an origin of replication functional in at least one organism, convenient restriction endonuclease sites, and a selectable marker for the host cell. Vectors according to the invention include expression vectors, replication vectors, probe generation vectors, and sequencing vectors. A host cell according to the invention can be a prokaryotic or eukaryotic cell and can be a unicellular organism or part of a multicellular organism.

The present invention further provides recombinant constructs comprising a nucleic acid having any of the nucleotide sequences of SEQ ID NO:1-1009 or a fragment thereof or any other polynucleotides of the invention. In one embodiment, the recombinant constructs of the present invention comprise a vector, such as a plasmid or viral vector, into which a nucleic acid having any of the nucleotide sequences of SEQ ID NO:1-1009 or a fragment thereof is inserted, in a forward or reverse orientation. In the case of a vector comprising one of the ORFs of the present invention, the vector may further comprise regulatory sequences, including for example, a promoter, operably linked to the ORF. Large numbers of suitable vectors and promoters are

known to those of skill in the art and are commercially available for generating the recombinant constructs of the present invention. The following vectors are provided by way of example.

Bacterial: pBs, phagescript, PsiX174, pBluescript SK, pBs KS, pNH8a, pNH16a, pNH18a, pNH46a (Stratagene); pTrc99A, pKK223-3, pKK233-3, pDR540, pRIT5 (Pharmacia).

5 Eukaryotic: pWLneo, pSV2cat, pOG44, PXTI, pSG (Stratagene) pSVK3, pBPV, pMSG, pSVL (Pharmacia).

The isolated polynucleotide of the invention may be operably linked to an expression control sequence such as the pMT2 or pED expression vectors disclosed in Kaufman et al., *Nucleic Acids Res.* 19, 4485-4490 (1991), in order to produce the protein recombinantly. Many 10 suitable expression control sequences are known in the art. General methods of expressing recombinant proteins are also known and are exemplified in R. Kaufman, *Methods in Enzymology* 185, 537-566 (1990). As defined herein "operably linked" means that the isolated polynucleotide of the invention and an expression control sequence are situated within a vector or cell in such a way that the protein is expressed by a host cell which has been transformed 15 (transfected) with the ligated polynucleotide/expression control sequence.

Promoter regions can be selected from any desired gene using CAT (chloramphenicol transferase) vectors or other vectors with selectable markers. Two appropriate vectors are pKK232-8 and pCM7. Particular named bacterial promoters include lacI, lacZ, T3, T7, gpt, lambda PR, and trc. Eukaryotic promoters include CMV immediate early, HSV thymidine 20 kinase, early and late SV40, LTRs from retrovirus, and mouse metallothionein-I. Selection of the appropriate vector and promoter is well within the level of ordinary skill in the art.

Generally, recombinant expression vectors will include origins of replication and selectable markers permitting transformation of the host cell, e.g., the ampicillin resistance gene of *E. coli* and *S. cerevisiae* TRP1 gene, and a promoter derived from a highly-expressed gene to direct 25 transcription of a downstream structural sequence. Such promoters can be derived from operons encoding glycolytic enzymes such as 3-phosphoglycerate kinase (PGK), a-factor, acid phosphatase, or heat shock proteins, among others. The heterologous structural sequence is assembled in appropriate phase with translation initiation and termination sequences, and preferably, a leader sequence capable of directing secretion of translated protein into the 30 periplasmic space or extracellular medium. Optionally, the heterologous sequence can encode a fusion protein including an amino terminal identification peptide imparting desired characteristics, e.g., stabilization or simplified purification of expressed recombinant product.

Useful expression vectors for bacterial use are constructed by inserting a structural DNA sequence encoding a desired protein together with suitable translation initiation and termination 35 signals in operable reading phase with a functional promoter. The vector will comprise one or

more phenotypic selectable markers and an origin of replication to ensure maintenance of the vector and to, if desirable, provide amplification within the host. Suitable prokaryotic hosts for transformation include *E. coli*, *Bacillus subtilis*, *Salmonella typhimurium* and various species within the genera *Pseudomonas*, *Streptomyces*, and *Staphylococcus*, although others may also be 5 employed as a matter of choice.

As a representative but non-limiting example, useful expression vectors for bacterial use can comprise a selectable marker and bacterial origin of replication derived from commercially available plasmids comprising genetic elements of the well known cloning vector pBR322 (ATCC 37017). Such commercial vectors include, for example, pKK223-3 (Pharmacia Fine 10 Chemicals, Uppsala, Sweden) and GEM 1 (Promega Biotech, Madison, WI, USA). These pBR322 "backbone" sections are combined with an appropriate promoter and the structural sequence to be expressed. Following transformation of a suitable host strain and growth of the host strain to an appropriate cell density, the selected promoter is induced or derepressed by appropriate means (e.g., temperature shift or chemical induction) and cells are cultured for an 15 additional period. Cells are typically harvested by centrifugation, disrupted by physical or chemical means, and the resulting crude extract retained for further purification.

Polynucleotides of the invention can also be used to induce immune responses. For example, as described in Fan et al., *Nat. Biotech.* 17:870-872 (1999), incorporated herein by reference, nucleic acid sequences encoding a polypeptide may be used to generate antibodies 20 against the encoded polypeptide following topical administration of naked plasmid DNA or following injection, and preferably intramuscular injection of the DNA. The nucleic acid sequences are preferably inserted in a recombinant expression vector and may be in the form of naked DNA.

#### 25 4.3 ANTISENSE

Another aspect of the invention pertains to isolated antisense nucleic acid molecules that are hybridizable to or complementary to the nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:1-1009, or fragments, analogs or derivatives thereof. An "antisense" 30 nucleic acid comprises a nucleotide sequence that is complementary to a "sense" nucleic acid encoding a protein, e.g., complementary to the coding strand of a double-stranded cDNA molecule or complementary to an mRNA sequence. In specific aspects, antisense nucleic acid molecules are provided that comprise a sequence complementary to at least about 10, 25, 50, 100, 250 or 500 nucleotides or an entire coding strand, or to only a portion thereof. Nucleic acid molecules encoding fragments, homologs, derivatives and analogs of a protein of any of SEQ ID

NO:1010-2018 or antisense nucleic acids complementary to a nucleic acid sequence of SEQ ID NO:1-1009 are additionally provided.

In one embodiment, an antisense nucleic acid molecule is antisense to a "coding region" of the coding strand of a nucleotide sequence of the invention. The term "coding region" refers to the region of the nucleotide sequence comprising codons which are translated into amino acid residues. In another embodiment, the antisense nucleic acid molecule is antisense to a "noncoding region" of the coding strand of a nucleotide sequence of the invention. The term "noncoding region" refers to 5' and 3' sequences which flank the coding region that are not translated into amino acids (*i.e.*, also referred to as 5' and 3' untranslated regions).

Given the coding strand sequences encoding a nucleic acid disclosed herein (*e.g.*, SEQ ID NO:1-1009), antisense nucleic acids of the invention can be designed according to the rules of Watson and Crick or Hoogsteen base pairing. The antisense nucleic acid molecule can be complementary to the entire coding region of a mRNA, but more preferably is an oligonucleotide that is antisense to only a portion of the coding or noncoding region of a mRNA. For example, the antisense oligonucleotide can be complementary to the region surrounding the translation start site of a mRNA. An antisense oligonucleotide can be, for example, about 5, 10, 15, 20, 25, 30, 35, 40, 45 or 50 nucleotides in length. An antisense nucleic acid of the invention can be constructed using chemical synthesis or enzymatic ligation reactions using procedures known in the art. For example, an antisense nucleic acid (*e.g.*, an antisense oligonucleotide) can be chemically synthesized using naturally occurring nucleotides or variously modified nucleotides designed to increase the biological stability of the molecules or to increase the physical stability of the duplex formed between the antisense and sense nucleic acids, *e.g.*, phosphorothioate derivatives and acridine substituted nucleotides can be used.

Examples of modified nucleotides that can be used to generate the antisense nucleic acid include: 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xanthine, 4-acetylcytosine, 5-(carboxyhydroxymethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiacytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine. Alternatively, the

antisense nucleic acid can be produced biologically using an expression vector into which a nucleic acid has been subcloned in an antisense orientation (*i.e.*, RNA transcribed from the inserted nucleic acid will be of an antisense orientation to a target nucleic acid of interest, described further in the following subsection).

5 The antisense nucleic acid molecules of the invention are typically administered to a subject or generated *in situ* such that they hybridize with or bind to cellular mRNA and/or genomic DNA encoding a protein according to the invention to thereby inhibit expression of the protein, *e.g.*, by inhibiting transcription and/or translation. The hybridization can be by conventional nucleotide complementarity to form a stable duplex, or, for example, in the case of  
10 an antisense nucleic acid molecule that binds to DNA duplexes, through specific interactions in the major groove of the double helix. An example of a route of administration of antisense nucleic acid molecules of the invention includes direct injection at a tissue site. Alternatively, antisense nucleic acid molecules can be modified to target selected cells and then administered  
15 systemically. For example, for systemic administration, antisense molecules can be modified such that they specifically bind to receptors or antigens expressed on a selected cell surface, *e.g.*, by linking the antisense nucleic acid molecules to peptides or antibodies that bind to cell surface receptors or antigens. The antisense nucleic acid molecules can also be delivered to cells using the vectors described herein. To achieve sufficient intracellular concentrations of antisense molecules, vector constructs in which the antisense nucleic acid molecule is placed under the  
20 control of a strong pol II or pol III promoter are preferred.

In yet another embodiment, the antisense nucleic acid molecule of the invention is an  $\alpha$ -anomeric nucleic acid molecule. An  $\alpha$ -anomeric nucleic acid molecule forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual  $\beta$ -units, the strands run parallel to each other (Gaultier *et al.* (1987) *Nucleic Acids Res* 15: 6625-6641). The  
25 antisense nucleic acid molecule can also comprise a 2'-*o*-methylribonucleotide (Inoue *et al.* (1987) *Nucleic Acids Res* 15: 6131-6148) or a chimeric RNA -DNA analogue (Inoue *et al.* (1987) *FEBS Lett* 215: 327-330).

#### 4.4 RIBOZYMES AND PNA MOIETIES

30 In still another embodiment, an antisense nucleic acid of the invention is a ribozyme. Ribozymes are catalytic RNA molecules with ribonuclease activity that are capable of cleaving a single-stranded nucleic acid, such as a mRNA, to which they have a complementary region. Thus, ribozymes (*e.g.*, hammerhead ribozymes (described in Haselhoff and Gerlach (1988)  
35 *Nature* 334:585-591)) can be used to catalytically cleave a mRNA transcripts to thereby inhibit translation of a mRNA. A ribozyme having specificity for a nucleic acid of the invention can be

designed based upon the nucleotide sequence of a DNA disclosed herein (*i.e.*, SEQ ID NO:1-1009). For example, a derivative of a Tetrahymena L-19 IVS RNA can be constructed in which the nucleotide sequence of the active site is complementary to the nucleotide sequence to be cleaved in a SECX-encoding mRNA. See, *e.g.*, Cech *et al.* U.S. Pat. No. 4,987,071; and Cech *et al.* U.S. Pat. No. 5,116,742. Alternatively, SECX mRNA can be used to select a catalytic RNA having a specific ribonuclease activity from a pool of RNA molecules. See, *e.g.*, Bartel *et al.*, (1993) *Science* 261:1411-1418.

10 Alternatively, gene expression can be inhibited by targeting nucleotide sequences complementary to the regulatory region (*e.g.*, promoter and/or enhancers) to form triple helical structures that prevent transcription of the gene in target cells. See generally, Helene. (1991) *Anticancer Drug Des.* 6: 569-84; Helene. *et al.* (1992) *Ann. N.Y. Acad. Sci.* 660:27-36; and Maher (1992) *Bioassays* 14: 807-15.

15 In various embodiments, the nucleic acids of the invention can be modified at the base moiety, sugar moiety or phosphate backbone to improve, *e.g.*, the stability, hybridization, or solubility of the molecule. For example, the deoxyribose phosphate backbone of the nucleic acids can be modified to generate peptide nucleic acids (see Hyrup *et al.* (1996) *Bioorg Med Chem* 4: 5-23). As used herein, the terms "peptide nucleic acids" or "PNAs" refer to nucleic acid mimics, *e.g.*, DNA mimics, in which the deoxyribose phosphate backbone is replaced by a pseudopeptide backbone and only the four natural nucleobases are retained. The neutral 20 backbone of PNAs has been shown to allow for specific hybridization to DNA and RNA under conditions of low ionic strength. The synthesis of PNA oligomers can be performed using standard solid phase peptide synthesis protocols as described in Hyrup *et al.* (1996) above; Perry-O'Keefe *et al.* (1996) *PNAS* 93: 14670-675.

25 PNAs of the invention can be used in therapeutic and diagnostic applications. For example, PNAs can be used as antisense or antigene agents for sequence-specific modulation of gene expression by, *e.g.*, inducing transcription or translation arrest or inhibiting replication. PNAs of the invention can also be used, *e.g.*, in the analysis of single base pair mutations in a gene by, *e.g.*, PNA directed PCR clamping; as artificial restriction enzymes when used in combination with other enzymes, *e.g.*, S1 nucleases (Hyrup B. (1996) above); or as probes or 30 primers for DNA sequence and hybridization (Hyrup *et al.* (1996), above; Perry-O'Keefe (1996), above).

35 In another embodiment, PNAs of the invention can be modified, *e.g.*, to enhance their stability or cellular uptake, by attaching lipophilic or other helper groups to PNA, by the formation of PNA-DNA chimeras, or by the use of liposomes or other techniques of drug delivery known in the art. For example, PNA-DNA chimeras can be generated that may

combine the advantageous properties of PNA and DNA. Such chimeras allow DNA recognition enzymes, *e.g.*, RNase H and DNA polymerases, to interact with the DNA portion while the PNA portion would provide high binding affinity and specificity. PNA-DNA chimeras can be linked using linkers of appropriate lengths selected in terms of base stacking, number of bonds between the nucleobases, and orientation (Hyrup (1996) above). The synthesis of PNA-DNA chimeras can be performed as described in Hyrup (1996) above and Finn *et al.* (1996) *Nucl Acids Res* 24: 3357-63. For example, a DNA chain can be synthesized on a solid support using standard phosphoramidite coupling chemistry, and modified nucleoside analogs, *e.g.*, 5'-(4-methoxytrityl)amino-5'-deoxy-thymidine phosphoramidite, can be used between the PNA and the 5' end of DNA (Mag *et al.* (1989) *Nucl Acid Res* 17: 5973-88). PNA monomers are then coupled in a stepwise manner to produce a chimeric molecule with a 5' PNA segment and a 3' DNA segment (Finn *et al.* (1996) above). Alternatively, chimeric molecules can be synthesized with a 5' DNA segment and a 3' PNA segment. See, Petersen *et al.* (1975) *Bioorg Med Chem Lett* 5: 1119-1124.

15 In other embodiments, the oligonucleotide may include other appended groups such as peptides (*e.g.*, for targeting host cell receptors *in vivo*), or agents facilitating transport across the cell membrane (see, *e.g.*, Letsinger *et al.*, 1989, *Proc. Natl. Acad. Sci. U.S.A.* 86:6553-6556; Lemaitre *et al.*, 1987, *Proc. Natl. Acad. Sci.* 84:648-652; PCT Publication No. W088/09810) or the blood-brain barrier (see, *e.g.*, PCT Publication No. W089/10134). In addition, 20 oligonucleotides can be modified with hybridization triggered cleavage agents (See, *e.g.*, Krol *et al.*, 1988, *BioTechniques* 6:958-976) or intercalating agents. (See, *e.g.*, Zon, 1988, *Pharm. Res.* 5: 539-549). To this end, the oligonucleotide may be conjugated to another molecule, *e.g.*, a peptide, a hybridization triggered cross-linking agent, a transport agent, a hybridization-triggered cleavage agent, etc.

25

#### 4.5 HOSTS

The present invention further provides host cells genetically engineered to contain the polynucleotides of the invention. For example, such host cells may contain nucleic acids of the invention introduced into the host cell using known transformation, transfection or infection 30 methods. The present invention still further provides host cells genetically engineered to express the polynucleotides of the invention, wherein such polynucleotides are in operative association with a regulatory sequence heterologous to the host cell which drives expression of the polynucleotides in the cell.

Knowledge of nucleic acid sequences allows for modification of cells to permit, or 35 increase, expression of endogenous polypeptide. Cells can be modified (*e.g.*, by homologous

recombination) to provide increased polypeptide expression by replacing, in whole or in part, the naturally occurring promoter with all or part of a heterologous promoter so that the cells express the polypeptide at higher levels. The heterologous promoter is inserted in such a manner that it is operatively linked to the encoding sequences. See, for example, PCT International Publication 5 No. WO94/12650, PCT International Publication No. WO92/20808, and PCT International Publication No. WO91/09955. It is also contemplated that, in addition to heterologous promoter DNA, amplifiable marker DNA (e.g., ada, dhfr, and the multifunctional CAD gene which encodes carbamyl phosphate synthase, aspartate transcarbamylase, and dihydroorotate) and/or 10 intron DNA may be inserted along with the heterologous promoter DNA. If linked to the coding sequence, amplification of the marker DNA by standard selection methods results in co-amplification of the desired protein coding sequences in the cells.

The host cell can be a higher eukaryotic host cell, such as a mammalian cell, a lower eukaryotic host cell, such as a yeast cell, or the host cell can be a prokaryotic cell, such as a bacterial cell. Introduction of the recombinant construct into the host cell can be effected by 15 calcium phosphate transfection, DEAE, dextran mediated transfection, or electroporation (Davis, L. et al., *Basic Methods in Molecular Biology* (1986)). The host cells containing one of the polynucleotides of the invention, can be used in conventional manners to produce the gene product encoded by the isolated fragment (in the case of an ORF) or can be used to produce a heterologous protein under the control of the EMF.

20 Any host/vector system can be used to express one or more of the ORFs of the present invention. These include, but are not limited to, eukaryotic hosts such as HeLa cells, Cv-1 cell, COS cells, 293 cells, and Sf9 cells, as well as prokaryotic host such as *E. coli* and *B. subtilis*. The most preferred cells are those which do not normally express the particular polypeptide or 25 protein or which expresses the polypeptide or protein at low natural level. Mature proteins can be expressed in mammalian cells, yeast, bacteria, or other cells under the control of appropriate promoters. Cell-free translation systems can also be employed to produce such proteins using RNAs derived from the DNA constructs of the present invention. Appropriate cloning and expression vectors for use with prokaryotic and eukaryotic hosts are described by Sambrook, et al., in *Molecular Cloning: A Laboratory Manual*, Second Edition, Cold Spring Harbor, New 30 York (1989), the disclosure of which is hereby incorporated by reference.

Various mammalian cell culture systems can also be employed to express recombinant protein. Examples of mammalian expression systems include the COS-7 lines of monkey kidney fibroblasts, described by Gluzman, *Cell* 23:175 (1981). Other cell lines capable of expressing a compatible vector are, for example, the C127, monkey COS cells, Chinese Hamster Ovary 35 (CHO) cells, human kidney 293 cells, human epidermal A431 cells, human Colo205 cells, 3T3

cells, CV-1 cells, other transformed primate cell lines, normal diploid cells, cell strains derived from *in vitro* culture of primary tissue, primary explants, HeLa cells, mouse L cells, BHK, HL-60, U937, HaK or Jurkat cells. Mammalian expression vectors will comprise an origin of replication, a suitable promoter and also any necessary ribosome binding sites, polyadenylation site, splice donor and acceptor sites, transcriptional termination sequences, and 5' flanking nontranscribed sequences. DNA sequences derived from the SV40 viral genome, for example, SV40 origin, early promoter, enhancer, splice, and polyadenylation sites may be used to provide the required nontranscribed genetic elements. Recombinant polypeptides and proteins produced in bacterial culture are usually isolated by initial extraction from cell pellets, followed by one or more salting-out, aqueous ion exchange or size exclusion chromatography steps. Protein refolding steps can be used, as necessary, in completing configuration of the mature protein. Finally, high performance liquid chromatography (HPLC) can be employed for final purification steps. Microbial cells employed in expression of proteins can be disrupted by any convenient method, including freeze-thaw cycling, sonication, mechanical disruption, or use of cell lysing agents.

Alternatively, it may be possible to produce the protein in lower eukaryotes such as yeast or insects or in prokaryotes such as bacteria. Potentially suitable yeast strains include *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, *Kluyveromyces* strains, *Candida*, or any yeast strain capable of expressing heterologous proteins. Potentially suitable bacterial strains include *Escherichia coli*, *Bacillus subtilis*, *Salmonella typhimurium*, or any bacterial strain capable of expressing heterologous proteins. If the protein is made in yeast or bacteria, it may be necessary to modify the protein produced therein, for example by phosphorylation or glycosylation of the appropriate sites, in order to obtain the functional protein. Such covalent attachments may be accomplished using known chemical or enzymatic methods.

In another embodiment of the present invention, cells and tissues may be engineered to express an endogenous gene comprising the polynucleotides of the invention under the control of inducible regulatory elements, in which case the regulatory sequences of the endogenous gene may be replaced by homologous recombination. As described herein, gene targeting can be used to replace a gene's existing regulatory region with a regulatory sequence isolated from a different gene or a novel regulatory sequence synthesized by genetic engineering methods. Such regulatory sequences may be comprised of promoters, enhancers, scaffold-attachment regions, negative regulatory elements, transcriptional initiation sites, regulatory protein binding sites or combinations of said sequences. Alternatively, sequences which affect the structure or stability of the RNA or protein produced may be replaced, removed, added, or otherwise modified by targeting. These sequence include polyadenylation signals, mRNA stability elements, splice

sites, leader sequences for enhancing or modifying transport or secretion properties of the protein, or other sequences which alter or improve the function or stability of protein or RNA molecules.

The targeting event may be a simple insertion of the regulatory sequence, placing the 5 gene under the control of the new regulatory sequence, *e.g.*, inserting a new promoter or enhancer or both upstream of a gene. Alternatively, the targeting event may be a simple deletion of a regulatory element, such as the deletion of a tissue-specific negative regulatory element. Alternatively, the targeting event may replace an existing element; for example, a tissue-specific 10 enhancer can be replaced by an enhancer that has broader or different cell-type specificity than the naturally occurring elements. Here, the naturally occurring sequences are deleted and new sequences are added. In all cases, the identification of the targeting event may be facilitated by 15 the use of one or more selectable marker genes that are contiguous with the targeting DNA, allowing for the selection of cells in which the exogenous DNA has integrated into the host cell genome. The identification of the targeting event may also be facilitated by the use of one or more marker genes exhibiting the property of negative selection, such that the negatively 20 selectable marker is linked to the exogenous DNA, but configured such that the negatively selectable marker flanks the targeting sequence, and such that a correct homologous recombination event with sequences in the host cell genome does not result in the stable integration of the negatively selectable marker. Markers useful for this purpose include the Herpes Simplex Virus thymidine kinase (TK) gene or the bacterial xanthine-guanine phosphoribosyl-transferase (gpt) gene.

The gene targeting or gene activation techniques which can be used in accordance with this aspect of the invention are more particularly described in U.S. Patent No. 5,272,071 to Chappel; U.S. Patent No. 5,578,461 to Sherwin et al.; International Application No. 25 PCT/US92/09627 (WO93/09222) by Selden et al.; and International Application No. PCT/US90/06436 (WO91/06667) by Skoultchi et al., each of which is incorporated by reference herein in its entirety.

#### 4.6 POLYPEPTIDES OF THE INVENTION

30 The isolated polypeptides of the invention include, but are not limited to, a polypeptide comprising: the amino acid sequences set forth as any one of SEQ ID NO:1010-2018 or an amino acid sequence encoded by any one of the nucleotide sequences SEQ ID NO:1-1009 or the corresponding full length or mature protein. Polypeptides of the invention also include polypeptides preferably with biological or immunological activity that are encoded by: (a) a 35 polynucleotide having any one of the nucleotide sequences set forth in SEQ ID NO:1-1009 or (b)

polynucleotides encoding any one of the amino acid sequences set forth as SEQ ID NO:1010-2018 or (c) polynucleotides that hybridize to the complement of the polynucleotides of either (a) or (b) under stringent hybridization conditions. The invention also provides biologically active or immunologically active variants of any of the amino acid sequences set forth as SEQ ID NO:1010-2018 or the corresponding full length or mature protein; and "substantial equivalents" thereof (e.g., with at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, 86%, 87%, 88%, 89%, at least about 90%, 91%, 92%, 93%, 94%, typically at least about 95%, 96%, 97%, more typically at least about 98%, or most typically at least about 99% amino acid identity) that retain biological activity. Polypeptides encoded by allelic variants may have a similar, increased, or decreased activity compared to polypeptides comprising SEQ ID NO:1010-2018.

Fragments of the proteins of the present invention which are capable of exhibiting biological activity are also encompassed by the present invention. Fragments of the protein may be in linear form or they may be cyclized using known methods, for example, as described in H. 15 U. Saragovi, et al., *Bio/Technology* 10, 773-778 (1992) and in R. S. McDowell, et al., *J. Amer. Chem. Soc.* 114, 9245-9253 (1992), both of which are incorporated herein by reference. Such fragments may be fused to carrier molecules such as immunoglobulins for many purposes, including increasing the valency of protein binding sites.

The present invention also provides both full-length and mature forms (for example, 20 without a signal sequence or precursor sequence) of the disclosed proteins. The protein coding sequence is identified in the sequence listing by translation of the disclosed nucleotide sequences. The mature form of such protein may be obtained by expression of a full-length polynucleotide in a suitable mammalian cell or other host cell. The sequence of the mature form of the protein is also determinable from the amino acid sequence of the full-length form. Where 25 proteins of the present invention are membrane bound, soluble forms of the proteins are also provided. In such forms, part or all of the regions causing the proteins to be membrane bound are deleted so that the proteins are fully secreted from the cell in which they are expressed.

Protein compositions of the present invention may further comprise an acceptable carrier, such as a hydrophilic, e.g., pharmaceutically acceptable, carrier.

30 The present invention further provides isolated polypeptides encoded by the nucleic acid fragments of the present invention or by degenerate variants of the nucleic acid fragments of the present invention. By "degenerate variant" is intended nucleotide fragments which differ from a nucleic acid fragment of the present invention (e.g., an ORF) by nucleotide sequence but, due to the degeneracy of the genetic code, encode an identical polypeptide sequence. Preferred nucleic acid fragments of the present invention are the ORFs that encode proteins.

A variety of methodologies known in the art can be utilized to obtain any one of the isolated polypeptides or proteins of the present invention. At the simplest level, the amino acid sequence can be synthesized using commercially available peptide synthesizers. The synthetically-constructed protein sequences, by virtue of sharing primary, secondary or tertiary structural and/or conformational characteristics with proteins may possess biological properties in common therewith, including protein activity. This technique is particularly useful in producing small peptides and fragments of larger polypeptides. Fragments are useful, for example, in generating antibodies against the native polypeptide. Thus, they may be employed as biologically active or immunological substitutes for natural, purified proteins in screening of therapeutic compounds and in immunological processes for the development of antibodies.

The polypeptides and proteins of the present invention can alternatively be purified from cells which have been altered to express the desired polypeptide or protein. As used herein, a cell is said to be altered to express a desired polypeptide or protein when the cell, through genetic manipulation, is made to produce a polypeptide or protein which it normally does not produce or which the cell normally produces at a lower level. One skilled in the art can readily adapt procedures for introducing and expressing either recombinant or synthetic sequences into eukaryotic or prokaryotic cells in order to generate a cell which produces one of the polypeptides or proteins of the present invention.

The invention also relates to methods for producing a polypeptide comprising growing a culture of host cells of the invention in a suitable culture medium, and purifying the protein from the cells or the culture in which the cells are grown. For example, the methods of the invention include a process for producing a polypeptide in which a host cell containing a suitable expression vector that includes a polynucleotide of the invention is cultured under conditions that allow expression of the encoded polypeptide. The polypeptide can be recovered from the culture, conveniently from the culture medium, or from a lysate prepared from the host cells and further purified. Preferred embodiments include those in which the protein produced by such process is a full length or mature form of the protein.

In an alternative method, the polypeptide or protein is purified from bacterial cells which naturally produce the polypeptide or protein. One skilled in the art can readily follow known methods for isolating polypeptides and proteins in order to obtain one of the isolated polypeptides or proteins of the present invention. These include, but are not limited to, immunochromatography, HPLC, size-exclusion chromatography, ion-exchange chromatography, and immuno-affinity chromatography. See, e.g., Scopes, *Protein Purification: Principles and Practice*, Springer-Verlag (1994); Sambrook, et al., in *Molecular Cloning: A Laboratory Manual*; Ausubel et al., *Current Protocols in Molecular Biology*. Polypeptide fragments that

retain biological/immunological activity include fragments comprising greater than about 100 amino acids, or greater than about 200 amino acids, and fragments that encode specific protein domains.

The purified polypeptides can be used in *in vitro* binding assays which are well known in the art to identify molecules which bind to the polypeptides. These molecules include but are not limited to, for *e.g.*, small molecules, molecules from combinatorial libraries, antibodies or other proteins. The molecules identified in the binding assay are then tested for antagonist or agonist activity in *in vivo* tissue culture or animal models that are well known in the art. In brief, the molecules are titrated into a plurality of cell cultures or animals and then tested for either cell/animal death or prolonged survival of the animal/cells.

In addition, the peptides of the invention or molecules capable of binding to the peptides may be complexed with toxins, *e.g.*, ricin or cholera, or with other compounds that are toxic to cells. The toxin-binding molecule complex is then targeted to a tumor or other cell by the specificity of the binding molecule for SEQ ID NO:1010-2018.

The protein of the invention may also be expressed as a product of transgenic animals, *e.g.*, as a component of the milk of transgenic cows, goats, pigs, or sheep which are characterized by somatic or germ cells containing a nucleotide sequence encoding the protein.

The proteins provided herein also include proteins characterized by amino acid sequences similar to those of purified proteins but into which modification are naturally provided or deliberately engineered. For example, modifications, in the peptide or DNA sequence, can be made by those skilled in the art using known techniques. Modifications of interest in the protein sequences may include the alteration, substitution, replacement, insertion or deletion of a selected amino acid residue in the coding sequence. For example, one or more of the cysteine residues may be deleted or replaced with another amino acid to alter the conformation of the molecule. Techniques for such alteration, substitution, replacement, insertion or deletion are well known to those skilled in the art (see, *e.g.*, U.S. Pat. No. 4,518,584). Preferably, such alteration, substitution, replacement, insertion or deletion retains the desired activity of the protein. Regions of the protein that are important for the protein function can be determined by various methods known in the art including the alanine-scanning method which involved systematic substitution of single or strings of amino acids with alanine, followed by testing the resulting alanine-containing variant for biological activity. This type of analysis determines the importance of the substituted amino acid(s) in biological activity. Regions of the protein that are important for protein function may be determined by the eMATRIX program.

Other fragments and derivatives of the sequences of proteins which would be expected to retain protein activity in whole or in part and are useful for screening or other immunological

methodologies may also be easily made by those skilled in the art given the disclosures herein. Such modifications are encompassed by the present invention.

The protein may also be produced by operably linking the isolated polynucleotide of the invention to suitable control sequences in one or more insect expression vectors, and employing 5 an insect expression system. Materials and methods for baculovirus/insect cell expression systems are commercially available in kit form from, *e.g.*, Invitrogen, San Diego, Calif., U.S.A. (the MaxBat<sup>TM</sup> kit), and such methods are well known in the art, as described in Summers and Smith, Texas Agricultural Experiment Station Bulletin No. 1555 (1987), incorporated herein by reference. As used herein, an insect cell capable of expressing a polynucleotide of the present 10 invention is "transformed."

The protein of the invention may be prepared by culturing transformed host cells under culture conditions suitable to express the recombinant protein. The resulting expressed protein may then be purified from such culture (*i.e.*, from culture medium or cell extracts) using known purification processes, such as gel filtration and ion exchange chromatography. The purification 15 of the protein may also include an affinity column containing agents which will bind to the protein; one or more column steps over such affinity resins as concanavalin A-agarose, heparin-toyopearl<sup>TM</sup> or Cibacrom blue 3GA Sepharose<sup>TM</sup>; one or more steps involving hydrophobic interaction chromatography using such resins as phenyl ether, butyl ether, or propyl ether; or immunoaffinity chromatography.

20 Alternatively, the protein of the invention may also be expressed in a form which will facilitate purification. For example, it may be expressed as a fusion protein, such as those of maltose binding protein (MBP), glutathione-S-transferase (GST) or thioredoxin (TRX), or as a His tag. Kits for expression and purification of such fusion proteins are commercially available from New England BioLab (Beverly, Mass.), Pharmacia (Piscataway, N.J.) and Invitrogen, 25 respectively. The protein can also be tagged with an epitope and subsequently purified by using a specific antibody directed to such epitope. One such epitope ("FLAG<sup>®</sup>") is commercially available from Kodak (New Haven, Conn.).

Finally, one or more reverse-phase high performance liquid chromatography (RP- HPLC) 30 steps employing hydrophobic RP-HPLC media, *e.g.*, silica gel having pendant methyl or other aliphatic groups, can be employed to further purify the protein. Some or all of the foregoing purification steps, in various combinations, can also be employed to provide a substantially homogeneous isolated recombinant protein. The protein thus purified is substantially free of other mammalian proteins and is defined in accordance with the present invention as an "isolated protein."

The polypeptides of the invention include analogs (variants). This embraces fragments, as well as peptides in which one or more amino acids has been deleted, inserted, or substituted. Also, analogs of the polypeptides of the invention embrace fusions of the polypeptides or modifications of the polypeptides of the invention, wherein the polypeptide or analog is fused to another moiety or moieties, *e.g.*, targeting moiety or another therapeutic agent. Such analogs may exhibit improved properties such as activity and/or stability. Examples of moieties which may be fused to the polypeptide or an analog include, for example, targeting moieties which provide for the delivery of polypeptide to pancreatic cells, *e.g.*, antibodies to pancreatic cells, antibodies to immune cells such as T-cells, monocytes, dendritic cells, granulocytes, etc., as well as receptor and ligands expressed on pancreatic or immune cells. Other moieties which may be fused to the polypeptide include therapeutic agents which are used for treatment, for example, immunosuppressive drugs such as cyclosporin, SK506, azathioprine, CD3 antibodies and steroids. Also, polypeptides may be fused to immune modulators, and other cytokines such as alpha or beta interferon.

15

#### **4.6.1 DETERMINING POLYPEPTIDE AND POLYNUCLEOTIDE IDENTITY AND SIMILARITY**

Preferred identity and/or similarity are designed to give the largest match between the sequences tested. Methods to determine identity and similarity are codified in computer programs including, but are not limited to, the GCG program package, including GAP (Devereux, J., et al., Nucleic Acids Research 12(1):387 (1984); Genetics Computer Group, University of Wisconsin, Madison, WI), BLASTP, BLASTN, BLASTX, FASTA (Altschul, S.F. et al., J. Molec. Biol. 215:403-410 (1990), PSI-BLAST (Altschul S.F. et al., Nucleic Acids Res. vol. 25, pp. 3389-3402, herein incorporated by reference), eMatrix software (Wu et al., J. Comp. Biol., Vol. 6, pp. 219-235 (1999), herein incorporated by reference), eMotif software (Nevill-Manning et al, ISMB-97, Vol. 4, pp. 202-209, herein incorporated by reference), pFam software (Sonnhammer et al., Nucleic Acids Res., Vol. 26(1), pp. 320-322 (1998), herein incorporated by reference) and the Kyte-Doolittle hydrophobicity prediction algorithm (J. Mol Biol, 157, pp. 105-31 (1982), incorporated herein by reference). The BLAST programs are publicly available from the National Center for Biotechnology Information (NCBI) and other sources (BLAST Manual, Altschul, S., et al. NCB NLM NIH Bethesda, MD 20894; Altschul, S., et al., J. Mol. Biol. 215:403-410 (1990)).

#### **4.7 CHIMERIC AND FUSION PROTEINS**

The invention also provides chimeric or fusion proteins. As used herein, a "chimeric protein" or "fusion protein" comprises a polypeptide of the invention operatively linked to

another polypeptide. Within a fusion protein the polypeptide according to the invention can correspond to all or a portion of a protein according to the invention. In one embodiment, a fusion protein comprises at least one biologically active portion of a protein according to the invention. In another embodiment, a fusion protein comprises at least two biologically active 5 portions of a protein according to the invention. Within the fusion protein, the term "operatively linked" is intended to indicate that the polypeptide according to the invention and the other polypeptide are fused in-frame to each other. The polypeptide can be fused to the N-terminus or C-terminus.

For example, in one embodiment a fusion protein comprises a polypeptide according to 10 the invention operably linked to the extracellular domain of a second protein.

In another embodiment, the fusion protein is a GST-fusion protein in which the polypeptide sequences of the invention are fused to the C-terminus of the GST (*i.e.*, glutathione S-transferase) sequences.

In another embodiment, the fusion protein is an immunoglobulin fusion protein in which 15 the polypeptide sequences according to the invention comprises one or more domains are fused to sequences derived from a member of the immunoglobulin protein family. The immunoglobulin fusion proteins of the invention can be incorporated into pharmaceutical compositions and administered to a subject to inhibit an interaction between a ligand and a protein of the invention on the surface of a cell, to thereby suppress signal transduction *in vivo*. 20 The immunoglobulin fusion proteins can be used to affect the bioavailability of a cognate ligand. Inhibition of the ligand/protein interaction may be useful therapeutically for both the treatment of proliferative and differentiative disorders, *e.g.*, cancer as well as modulating (*e.g.*, promoting or inhibiting) cell survival. Moreover, the immunoglobulin fusion proteins of the invention can be used as immunogens to produce antibodies in a subject, to purify ligands, and in screening assays 25 to identify molecules that inhibit the interaction of a polypeptide of the invention with a ligand.

A chimeric or fusion protein of the invention can be produced by standard recombinant DNA techniques. For example, DNA fragments coding for the different polypeptide sequences are ligated together in-frame in accordance with conventional techniques, *e.g.*, by employing 30 blunt-ended or stagger-ended termini for ligation, restriction enzyme digestion to provide for appropriate termini, filling-in of cohesive ends as appropriate, alkaline phosphatase treatment to avoid undesirable joining, and enzymatic ligation. In another embodiment, the fusion gene can be synthesized by conventional techniques including automated DNA synthesizers.

Alternatively, PCR amplification of gene fragments can be carried out using anchor primers that give rise to complementary overhangs between two consecutive gene fragments that can 35 subsequently be annealed and reamplified to generate a chimeric gene sequence (see, for

example, Ausubel et al. (eds.) *CURRENT PROTOCOLS IN MOLECULAR BIOLOGY*, John Wiley & Sons, 1992). Moreover, many expression vectors are commercially available that already encode a fusion moiety (e.g., a GST polypeptide). A nucleic acid encoding a polypeptide of the invention can be cloned into such an expression vector such that the fusion moiety is linked 5 in-frame to the protein of the invention.

#### 4.8 GENE THERAPY

Mutations in the polynucleotides of the invention gene may result in loss of normal function of the encoded protein. The invention thus provides gene therapy to restore normal 10 activity of the polypeptides of the invention; or to treat disease states involving polypeptides of the invention. Delivery of a functional gene encoding polypeptides of the invention to appropriate cells is effected *ex vivo*, *in situ*, or *in vivo* by use of vectors, and more particularly viral vectors (e.g., adenovirus, adeno-associated virus, or a retrovirus), or *ex vivo* by use of physical DNA transfer methods (e.g., liposomes or chemical treatments). See, for example, 15 Anderson, *Nature*, supplement to vol. 392, no. 6679, pp.25-20 (1998). For additional reviews of gene therapy technology see Friedmann, *Science*, 244: 1275-1281 (1989); Verma, *Scientific American*: 68-84 (1990); and Miller, *Nature*, 357: 455-460 (1992). Introduction of any one of 20 the nucleotides of the present invention or a gene encoding the polypeptides of the present invention can also be accomplished with extrachromosomal substrates (transient expression) or artificial chromosomes (stable expression). Cells may also be cultured *ex vivo* in the presence of 25 proteins of the present invention in order to proliferate or to produce a desired effect on or activity in such cells. Treated cells can then be introduced *in vivo* for therapeutic purposes. Alternatively, it is contemplated that in other human disease states, preventing the expression of or inhibiting the activity of polypeptides of the invention will be useful in treating the disease 25 states. It is contemplated that antisense therapy or gene therapy could be applied to negatively regulate the expression of polypeptides of the invention.

Other methods inhibiting expression of a protein include the introduction of antisense 30 molecules to the nucleic acids of the present invention, their complements, or their translated RNA sequences, by methods known in the art. Further, the polypeptides of the present invention can be inhibited by using targeted deletion methods, or the insertion of a negative regulatory element such as a silencer, which is tissue specific.

The present invention still further provides cells genetically engineered *in vivo* to express the polynucleotides of the invention, wherein such polynucleotides are in operative association with a regulatory sequence heterologous to the host cell which drives expression of the polynucleotides in

the cell. These methods can be used to increase or decrease the expression of the polynucleotides of the present invention.

Knowledge of DNA sequences provided by the invention allows for modification of cells to permit, increase, or decrease, expression of endogenous polypeptide. Cells can be modified (e.g., by 5 homologous recombination) to provide increased polypeptide expression by replacing, in whole or in part, the naturally occurring promoter with all or part of a heterologous promoter so that the cells express the protein at higher levels. The heterologous promoter is inserted in such a manner that it is operatively linked to the desired protein encoding sequences. See, for example, PCT International Publication No. WO 94/12650, PCT International Publication No. WO 92/20808, and PCT 10 International Publication No. WO 91/09955. It is also contemplated that, in addition to heterologous promoter DNA, amplifiable marker DNA (e.g., ada, dhfr, and the multifunctional CAD gene which encodes carbamyl phosphate synthase, aspartate transcarbamylase, and dihydroorotase) and/or intron DNA may be inserted along with the heterologous promoter DNA. If linked to the desired protein coding sequence, amplification of the marker DNA by standard selection methods results in 15 co-amplification of the desired protein coding sequences in the cells.

In another embodiment of the present invention, cells and tissues may be engineered to express an endogenous gene comprising the polynucleotides of the invention under the control of inducible regulatory elements, in which case the regulatory sequences of the endogenous gene may be replaced by homologous recombination. As described herein, gene targeting can be used to 20 replace a gene's existing regulatory region with a regulatory sequence isolated from a different gene or a novel regulatory sequence synthesized by genetic engineering methods. Such regulatory sequences may be comprised of promoters, enhancers, scaffold-attachment regions, negative regulatory elements, transcriptional initiation sites, regulatory protein binding sites or combinations of said sequences. Alternatively, sequences which affect the structure or stability of the RNA or 25 protein produced may be replaced, removed, added, or otherwise modified by targeting. These sequences include polyadenylation signals, mRNA stability elements, splice sites, leader sequences for enhancing or modifying transport or secretion properties of the protein, or other sequences which alter or improve the function or stability of protein or RNA molecules.

The targeting event may be a simple insertion of the regulatory sequence, placing the gene 30 under the control of the new regulatory sequence, e.g., inserting a new promoter or enhancer or both upstream of a gene. Alternatively, the targeting event may be a simple deletion of a regulatory element, such as the deletion of a tissue-specific negative regulatory element. Alternatively, the targeting event may replace an existing element; for example, a tissue-specific enhancer can be replaced by an enhancer that has broader or different cell-type specificity than the naturally 35 occurring elements. Here, the naturally occurring sequences are deleted and new sequences are

added. In all cases, the identification of the targeting event may be facilitated by the use of one or more selectable marker genes that are contiguous with the targeting DNA, allowing for the selection of cells in which the exogenous DNA has integrated into the cell genome. The identification of the targeting event may also be facilitated by the use of one or more marker genes exhibiting the 5 property of negative selection, such that the negatively selectable marker is linked to the exogenous DNA, but configured such that the negatively selectable marker flanks the targeting sequence, and such that a correct homologous recombination event with sequences in the host cell genome does not result in the stable integration of the negatively selectable marker. Markers useful for this purpose include the Herpes Simplex Virus thymidine kinase (TK) gene or the bacterial 10 xanthine-guanine phosphoribosyl-transferase (gpt) gene.

The gene targeting or gene activation techniques which can be used in accordance with this aspect of the invention are more particularly described in U.S. Patent No. 5,272,071 to Chappel; U.S. Patent No. 5,578,461 to Sherwin et al.; International Application No. PCT/US92/09627 (WO93/09222) by Selden et al.; and International Application No. PCT/US90/06436 15 (WO91/06667) by Skoultschi et al., each of which is incorporated by reference herein in its entirety.

#### 4.9 TRANSGENIC ANIMALS

In preferred methods to determine biological functions of the polypeptides of the invention in vivo, one or more genes provided by the invention are either over expressed or 20 inactivated in the germ line of animals using homologous recombination [Capecci, Science 244:1288-1292 (1989)]. Animals in which the gene is over expressed, under the regulatory control of exogenous or endogenous promoter elements, are known as transgenic animals. Animals in which an endogenous gene has been inactivated by homologous recombination are referred to as "knockout" animals. Knockout animals, preferably non-human mammals, can be 25 prepared as described in U.S. Patent No. 5,557,032, incorporated herein by reference. Transgenic animals are useful to determine the roles polypeptides of the invention play in biological processes, and preferably in disease states. Transgenic animals are useful as model systems to identify compounds that modulate lipid metabolism. Transgenic animals, preferably non-human mammals, are produced using methods as described in U.S. Patent No 5,489,743 and PCT 30 Publication No. WO94/28122, incorporated herein by reference.

Transgenic animals can be prepared wherein all or part of a promoter of the polynucleotides of the invention is either activated or inactivated to alter the level of expression of the polypeptides of the invention. Inactivation can be carried out using homologous recombination methods described above. Activation can be achieved by supplementing or even 35 replacing the homologous promoter to provide for increased protein expression. The homologous

promoter can be supplemented by insertion of one or more heterologous enhancer elements known to confer promoter activation in a particular tissue.

The polynucleotides of the present invention also make possible the development, through, *e.g.*, homologous recombination or knock out strategies, of animals that fail to express 5 polypeptides of the invention or that express a variant polypeptide. Such animals are useful as models for studying the *in vivo* activities of polypeptide as well as for studying modulators of the polypeptides of the invention.

In preferred methods to determine biological functions of the polypeptides of the invention *in vivo*, one or more genes provided by the invention are either over expressed or 10 inactivated in the germ line of animals using homologous recombination [Capecci, *Science* 244:1288-1292 (1989)]. Animals in which the gene is over expressed, under the regulatory control of exogenous or endogenous promoter elements, are known as transgenic animals. Animals in which an endogenous gene has been inactivated by homologous recombination are referred to as "knockout" animals. Knockout animals, preferably non-human mammals, can be 15 prepared as described in U.S. Patent No. 5,557,032, incorporated herein by reference. Transgenic animals are useful to determine the roles polypeptides of the invention play in biological processes, and preferably in disease states. Transgenic animals are useful as model systems to identify compounds that modulate lipid metabolism. Transgenic animals, preferably non-human mammals, are produced using methods as described in U.S. Patent No 5,489,743 and PCT 20 Publication No. WO94/28122, incorporated herein by reference.

Transgenic animals can be prepared wherein all or part of the polynucleotides of the invention promoter is either activated or inactivated to alter the level of expression of the polypeptides of the invention. Inactivation can be carried out using homologous recombination methods described above. Activation can be achieved by supplementing or even replacing the 25 homologous promoter to provide for increased protein expression. The homologous promoter can be supplemented by insertion of one or more heterologous enhancer elements known to confer promoter activation in a particular tissue.

#### **4.10 USES AND BIOLOGICAL ACTIVITY**

The polynucleotides and proteins of the present invention are expected to exhibit one or 30 more of the uses or biological activities (including those associated with assays cited herein) identified herein. Uses or activities described for proteins of the present invention may be provided by administration or use of such proteins or of polynucleotides encoding such proteins (such as, for example, in gene therapies or vectors suitable for introduction of DNA). The 35 mechanism underlying the particular condition or pathology will dictate whether the

polypeptides of the invention, the polynucleotides of the invention or modulators (activators or inhibitors) thereof would be beneficial to the subject in need of treatment. Thus, "therapeutic compositions of the invention" include compositions comprising isolated polynucleotides (including recombinant DNA molecules, cloned genes and degenerate variants thereof) or

5 polypeptides of the invention (including full length protein, mature protein and truncations or domains thereof), or compounds and other substances that modulate the overall activity of the target gene products, either at the level of target gene/protein expression or target protein activity. Such modulators include polypeptides, analogs, (variants), including fragments and fusion proteins, antibodies and other binding proteins; chemical compounds that directly or

10 indirectly activate or inhibit the polypeptides of the invention (identified, *e.g.*, via drug screening assays as described herein); antisense polynucleotides and polynucleotides suitable for triple helix formation; and in particular antibodies or other binding partners that specifically recognize one or more epitopes of the polypeptides of the invention.

The polypeptides of the present invention may likewise be involved in cellular activation

15 or in one of the other physiological pathways described herein.

#### 4.10.1 RESEARCH USES AND UTILITIES

The polynucleotides provided by the present invention can be used by the research community for various purposes. The polynucleotides can be used to express recombinant

20 protein for analysis, characterization or therapeutic use; as markers for tissues in which the corresponding protein is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in disease states); as molecular weight markers on gels; as chromosome markers or tags (when labeled) to identify chromosomes or to map related gene positions; to compare with endogenous DNA sequences in patients to identify potential genetic

25 disorders; as probes to hybridize and thus discover novel, related DNA sequences; as a source of information to derive PCR primers for genetic fingerprinting; as a probe to "subtract-out" known sequences in the process of discovering other novel polynucleotides; for selecting and making oligomers for attachment to a "gene chip" or other support, including for examination of expression patterns; to raise anti-protein antibodies using DNA immunization techniques; and as

30 an antigen to raise anti-DNA antibodies or elicit another immune response. Where the polynucleotide encodes a protein which binds or potentially binds to another protein (such as, for example, in a receptor-ligand interaction), the polynucleotide can also be used in interaction trap assays (such as, for example, that described in Gyuris et al., Cell 75:791-803 (1993)) to identify polynucleotides encoding the other protein with which binding occurs or to identify inhibitors of

35 the binding interaction.

The polypeptides provided by the present invention can similarly be used in assays to determine biological activity, including in a panel of multiple proteins for high-throughput screening; to raise antibodies or to elicit another immune response; as a reagent (including the labeled reagent) in assays designed to quantitatively determine levels of the protein (or its receptor) in biological fluids; as markers for tissues in which the corresponding polypeptide is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in a disease state); and, of course, to isolate correlative receptors or ligands. Proteins involved in these binding interactions can also be used to screen for peptide or small molecule inhibitors or agonists of the binding interaction.

Any or all of these research utilities are capable of being developed into reagent grade or kit format for commercialization as research products.

Methods for performing the uses listed above are well known to those skilled in the art. References disclosing such methods include without limitation "Molecular Cloning: A Laboratory Manual", 2d ed., Cold Spring Harbor Laboratory Press, Sambrook, J., E. F. Fritsch and T. Maniatis eds., 1989, and "Methods in Enzymology: Guide to Molecular Cloning Techniques", Academic Press, Berger, S. L. and A. R. Kimmel eds., 1987.

#### 4.10.2 NUTRITIONAL USES

Polynucleotides and polypeptides of the present invention can also be used as nutritional sources or supplements. Such uses include without limitation use as a protein or amino acid supplement, use as a carbon source, use as a nitrogen source and use as a source of carbohydrate. In such cases the polypeptide or polynucleotide of the invention can be added to the feed of a particular organism or can be administered as a separate solid or liquid preparation, such as in the form of powder, pills, solutions, suspensions or capsules. In the case of microorganisms, the polypeptide or polynucleotide of the invention can be added to the medium in or on which the microorganism is cultured.

#### 4.10.3 CYTOKINE AND CELL PROLIFERATION/DIFFERENTIATION ACTIVITY

A polypeptide of the present invention may exhibit activity relating to cytokine, cell proliferation (either inducing or inhibiting) or cell differentiation (either inducing or inhibiting) activity or may induce production of other cytokines in certain cell populations. A polynucleotide of the invention can encode a polypeptide exhibiting such attributes. Many protein factors discovered to date, including all known cytokines, have exhibited activity in one or more factor-dependent cell proliferation assays, and hence the assays serve as a convenient

confirmation of cytokine activity. The activity of therapeutic compositions of the present invention is evidenced by any one of a number of routine factor dependent cell proliferation assays for cell lines including, without limitation, 32D, DA2, DA1G, T10, B9, B9/11, BaF3, MC9/G, M+(preB M+), 2E8, RB5, DA1, 123, T1165, HT2, CTLL2, TF-1, Mo7e, CMK,

5 HUVEC, and Caco. Therapeutic compositions of the invention can be used in the following:

Assays for T-cell or thymocyte proliferation include without limitation those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, *In Vitro* assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Takai et al., *J. Immunol.* 137:3494-3500, 1986; Bertagnolli et al., *J. Immunol.* 145:1706-1712, 1990; Bertagnolli et al., *Cellular Immunology* 133:327-341, 1991; Bertagnolli, et al., *I. Immunol.* 149:3778-3783, 1992; Bowman et al., *I. Immunol.* 152:1756-1761, 1994.

Assays for cytokine production and/or proliferation of spleen cells, lymph node cells or thymocytes include, without limitation, those described in: Polyclonal T cell stimulation, Kruisbeek, A. M. and Shevach, E. M. In Current Protocols in Immunology. J. E. e.a. Coligan eds. Vol 1 pp. 3.12.1-3.12.14, John Wiley and Sons, Toronto. 1994; and Measurement of mouse and human interleukin- $\gamma$ , Schreiber, R. D. In Current Protocols in Immunology. J. E. e.a. Coligan eds. Vol 1 pp. 6.8.1-6.8.8, John Wiley and Sons, Toronto. 1994.

Assays for proliferation and differentiation of hematopoietic and lymphopoietic cells include, without limitation, those described in: Measurement of Human and Murine Interleukin 2 and Interleukin 4, Bottomly, K., Davis, L. S. and Lipsky, P. E. In Current Protocols in Immunology. J. E. e.a. Coligan eds. Vol 1 pp. 6.3.1-6.3.12, John Wiley and Sons, Toronto. 1991; deVries et al., *J. Exp. Med.* 173:1205-1211, 1991; Moreau et al., *Nature* 336:690-692, 1988; Greenberger et al., *Proc. Natl. Acad. Sci. U.S.A.* 80:2931-2938, 1983; Measurement of mouse and human interleukin 6--Nordan, R. In Current Protocols in Immunology. J. E. Coligan eds. Vol 1 pp. 6.6.1-6.6.5, John Wiley and Sons, Toronto. 1991; Smith et al., *Proc. Natl. Acad. Sci. U.S.A.* 83:1857-1861, 1986; Measurement of human Interleukin 11--Bennett, F., Giannotti, J., Clark, S. C. and Turner, K. J. In Current Protocols in Immunology. J. E. Coligan eds. Vol 1 pp. 6.15.1 John Wiley and Sons, Toronto. 1991; Measurement of mouse and human Interleukin 9--Ciarletta, A., Giannotti, J., Clark, S. C. and Turner, K. J. In Current Protocols in Immunology. J. E. Coligan eds. Vol 1 pp. 6.13.1, John Wiley and Sons, Toronto. 1991.

Assays for T-cell clone responses to antigens (which will identify, among others, proteins that affect APC-T cell interactions as well as direct T-cell effects by measuring proliferation and cytokine production) include, without limitation, those described in: Current Protocols in

35 Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W. Strober,

Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, *In Vitro* assays for Mouse Lymphocyte Function; Chapter 6, Cytokines and their cellular receptors; Chapter 7, Immunologic studies in Humans); Weinberger et al., Proc. Natl. Acad. Sci. USA 77:6091-6095, 1980; Weinberger et al., Eur. J. Immun. 11:405-411, 1981; Takai et al., J. Immunol. 5 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988.

#### 4.10.4 STEM CELL GROWTH FACTOR ACTIVITY

A polypeptide of the present invention may exhibit stem cell growth factor activity and be involved in the proliferation, differentiation and survival of pluripotent and totipotent stem 10 cells including primordial germ cells, embryonic stem cells, hematopoietic stem cells and/or germ line stem cells. Administration of the polypeptide of the invention to stem cells *in vivo* or *ex vivo* is expected to maintain and expand cell populations in a totipotential or pluripotential state which would be useful for re-engineering damaged or diseased tissues, transplantation, manufacture of bio-pharmaceuticals and the development of bio-sensors. The ability to produce 15 large quantities of human cells has important working applications for the production of human proteins which currently must be obtained from non-human sources or donors, implantation of cells to treat diseases such as Parkinson's, Alzheimer's and other neurodegenerative diseases; tissues for grafting such as bone marrow, skin, cartilage, tendons, bone, muscle (including 20 cardiac muscle), blood vessels, cornea, neural cells, gastrointestinal cells and others; and organs for transplantation such as kidney, liver, pancreas (including islet cells), heart and lung.

It is contemplated that multiple different exogenous growth factors and/or cytokines may be administered in combination with the polypeptide of the invention to achieve the desired effect, including any of the growth factors listed herein, other stem cell maintenance factors, and specifically including stem cell factor (SCF), leukemia inhibitory factor (LIF), Flt-3 ligand (Flt-25 3L), any of the interleukins, recombinant soluble IL-6 receptor fused to IL-6, macrophage inflammatory protein 1-alpha (MIP-1-alpha), G-CSF, GM-CSF, thrombopoietin (TPO), platelet factor 4 (PF-4), platelet-derived growth factor (PDGF), neural growth factors and basic fibroblast growth factor (bFGF).

Since totipotent stem cells can give rise to virtually any mature cell type, expansion of 30 these cells in culture will facilitate the production of large quantities of mature cells. Techniques for culturing stem cells are known in the art and administration of polypeptides of the invention, optionally with other growth factors and/or cytokines, is expected to enhance the survival and proliferation of the stem cell populations. This can be accomplished by direct administration of the polypeptide of the invention to the culture medium. Alternatively, stroma cells transfected 35 with a polynucleotide that encodes for the polypeptide of the invention can be used as a feeder

layer for the stem cell populations in culture or in vivo. Stromal support cells for feeder layers may include embryonic bone marrow fibroblasts, bone marrow stromal cells, fetal liver cells, or cultured embryonic fibroblasts (see U.S. Patent No. 5,690,926).

Stem cells themselves can be transfected with a polynucleotide of the invention to induce 5 autocrine expression of the polypeptide of the invention. This will allow for generation of undifferentiated totipotential/pluripotential stem cell lines that are useful as is or that can then be differentiated into the desired mature cell types. These stable cell lines can also serve as a source of undifferentiated totipotential/pluripotential mRNA to create cDNA libraries and templates for polymerase chain reaction experiments. These studies would allow for the isolation and 10 identification of differentially expressed genes in stem cell populations that regulate stem cell proliferation and/or maintenance.

Expansion and maintenance of totipotent stem cell populations will be useful in the treatment of many pathological conditions. For example, polypeptides of the present invention may be used to manipulate stem cells in culture to give rise to neuroepithelial cells that can be 15 used to augment or replace cells damaged by illness, autoimmune disease, accidental damage or genetic disorders. The polypeptide of the invention may be useful for inducing the proliferation of neural cells and for the regeneration of nerve and brain tissue, *i.e.* for the treatment of central and peripheral nervous system diseases and neuropathies, as well as mechanical and traumatic disorders which involve degeneration, death or trauma to neural cells or nerve tissue. In addition, 20 the expanded stem cell populations can also be genetically altered for gene therapy purposes and to decrease host rejection of replacement tissues after grafting or implantation.

Expression of the polypeptide of the invention and its effect on stem cells can also be manipulated to achieve controlled differentiation of the stem cells into more differentiated cell types. A broadly applicable method of obtaining pure populations of a specific differentiated 25 cell type from undifferentiated stem cell populations involves the use of a cell-type specific promoter driving a selectable marker. The selectable marker allows only cells of the desired type to survive. For example, stem cells can be induced to differentiate into cardiomyocytes (Wobus et al., Differentiation, 48: 173-182, (1991); Klug et al., J. Clin. Invest., 98(1): 216-224, (1998)) or skeletal muscle cells (Browder, L. W. In: *Principles of Tissue Engineering* eds. Lanza et al., 30 Academic Press (1997)). Alternatively, directed differentiation of stem cells can be accomplished by culturing the stem cells in the presence of a differentiation factor such as retinoic acid and an antagonist of the polypeptide of the invention which would inhibit the effects of endogenous stem cell factor activity and allow differentiation to proceed.

*In vitro* cultures of stem cells can be used to determine if the polypeptide of the invention 35 exhibits stem cell growth factor activity. Stem cells are isolated from any one of various cell

sources (including hematopoietic stem cells and embryonic stem cells) and cultured on a feeder layer, as described by Thompson et al. Proc. Natl. Acad. Sci, U.S.A., 92: 7844-7848 (1995), in the presence of the polypeptide of the invention alone or in combination with other growth factors or cytokines. The ability of the polypeptide of the invention to induce stem cells 5 proliferation is determined by colony formation on semi-solid support *e.g.* as described by Bernstein et al., Blood, 77: 2316-2321 (1991).

#### 4.10.5 HEMATOPOIESIS REGULATING ACTIVITY

A polypeptide of the present invention may be involved in regulation of hematopoiesis 10 and, consequently, in the treatment of myeloid or lymphoid cell disorders. Even marginal biological activity in support of colony forming cells or of factor-dependent cell lines indicates involvement in regulating hematopoiesis, *e.g.* in supporting the growth and proliferation of erythroid progenitor cells alone or in combination with other cytokines, thereby indicating utility, for example, in treating various anemias or for use in conjunction with 15 irradiation/chemotherapy to stimulate the production of erythroid precursors and/or erythroid cells; in supporting the growth and proliferation of myeloid cells such as granulocytes and monocytes/macrophages (*i.e.*, traditional CSF activity) useful, for example, in conjunction with chemotherapy to prevent or treat consequent myelo-suppression; in supporting the growth and proliferation of megakaryocytes and consequently of platelets thereby allowing prevention or 20 treatment of various platelet disorders such as thrombocytopenia, and generally for use in place of or complimentary to platelet transfusions; and/or in supporting the growth and proliferation of hematopoietic stem cells which are capable of maturing to any and all of the above-mentioned hematopoietic cells and therefore find therapeutic utility in various stem cell disorders (such as 25 those usually treated with transplantation, including, without limitation, aplastic anemia and paroxysmal nocturnal hemoglobinuria), as well as in repopulating the stem cell compartment post irradiation/chemotherapy, either *in-vivo* or *ex-vivo* (*i.e.*, in conjunction with bone marrow transplantation or with peripheral progenitor cell transplantation (homologous or heterologous)) as normal cells or genetically manipulated for gene therapy.

Therapeutic compositions of the invention can be used in the following:

30 Suitable assays for proliferation and differentiation of various hematopoietic lines are cited above.

Assays for embryonic stem cell differentiation (which will identify, among others, 35 proteins that influence embryonic differentiation hematopoiesis) include, without limitation, those described in: Johansson et al. Cellular Biology 15:141-151, 1995; Keller et al., Molecular and Cellular Biology 13:473-486, 1993; McClanahan et al., Blood 81:2903-2915, 1993.

Assays for stem cell survival and differentiation (which will identify, among others, proteins that regulate lympho-hematopoiesis) include, without limitation, those described in: Methylcellulose colony forming assays, Freshney, M. G. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 265-268, Wiley-Liss, Inc., New York, N.Y. 1994; Hirayama et al., 5 Proc. Natl. Acad. Sci. USA 89:5907-5911, 1992; Primitive hematopoietic colony forming cells with high proliferative potential, McNiece, I. K. and Briddell, R. A. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 23-39, Wiley-Liss, Inc., New York, N.Y. 1994; Neben et al., Experimental Hematology 22:353-359, 1994; Cobblestone area forming cell assay, Ploemacher, R. E. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 1-21, 10 Wiley-Liss, Inc., New York, N.Y. 1994; Long term bone marrow cultures in the presence of stromal cells, Spooncer, E., Dexter, M. and Allen, T. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 163-179, Wiley-Liss, Inc., New York, N.Y. 1994; Long term culture initiating cell assay, Sutherland, H. J. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 139-162, Wiley-Liss, Inc., New York, N.Y. 1994.

15

#### **4.10.6 TISSUE GROWTH ACTIVITY**

A polypeptide of the present invention also may be involved in bone, cartilage, tendon, ligament and/or nerve tissue growth or regeneration, as well as in wound healing and tissue repair and replacement, and in healing of burns, incisions and ulcers.

20

A polypeptide of the present invention which induces cartilage and/or bone growth in circumstances where bone is not normally formed, has application in the healing of bone fractures and cartilage damage or defects in humans and other animals. Compositions of a polypeptide, antibody, binding partner, or other modulator of the invention may have prophylactic use in closed as well as open fracture reduction and also in the improved fixation of 25 artificial joints. De novo bone formation induced by an osteogenic agent contributes to the repair of congenital, trauma induced, or oncologic resection induced craniofacial defects, and also is useful in cosmetic plastic surgery.

25

A polypeptide of this invention may also be involved in attracting bone-forming cells, stimulating growth of bone-forming cells, or inducing differentiation of progenitors of 30 bone-forming cells. Treatment of osteoporosis, osteoarthritis, bone degenerative disorders, or periodontal disease, such as through stimulation of bone and/or cartilage repair or by blocking inflammation or processes of tissue destruction (collagenase activity, osteoclast activity, etc.) mediated by inflammatory processes may also be possible using the composition of the invention.

Another category of tissue regeneration activity that may involve the polypeptide of the present invention is tendon/ligament formation. Induction of tendon/ligament-like tissue or other tissue formation in circumstances where such tissue is not normally formed, has application in the healing of tendon or ligament tears, deformities and other tendon or ligament 5 defects in humans and other animals. Such a preparation employing a tendon/ligament-like tissue inducing protein may have prophylactic use in preventing damage to tendon or ligament tissue, as well as use in the improved fixation of tendon or ligament to bone or other tissues, and in repairing defects to tendon or ligament tissue. De novo tendon/ligament-like tissue formation induced by a composition of the present invention contributes to the repair of congenital, trauma 10 induced, or other tendon or ligament defects of other origin, and is also useful in cosmetic plastic surgery for attachment or repair of tendons or ligaments. The compositions of the present invention may provide environment to attract tendon- or ligament-forming cells, stimulate growth of tendon- or ligament-forming cells, induce differentiation of progenitors of tendon- or ligament-forming cells, or induce growth of tendon/ligament cells or progenitors *ex vivo* for 15 return *in vivo* to effect tissue repair. The compositions of the invention may also be useful in the treatment of tendinitis, carpal tunnel syndrome and other tendon or ligament defects. The compositions may also include an appropriate matrix and/or sequestering agent as a carrier as is well known in the art.

The compositions of the present invention may also be useful for proliferation of neural 20 cells and for regeneration of nerve and brain tissue, *i.e.* for the treatment of central and peripheral nervous system diseases and neuropathies, as well as mechanical and traumatic disorders, which involve degeneration, death or trauma to neural cells or nerve tissue. More specifically, a composition may be used in the treatment of diseases of the peripheral nervous system, such as peripheral nerve injuries, peripheral neuropathy and localized neuropathies, and central nervous 25 system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager syndrome. Further conditions which may be treated in accordance with the present invention include mechanical and traumatic disorders, such as spinal cord disorders, head trauma and cerebrovascular diseases such as stroke. Peripheral neuropathies resulting from chemotherapy or other medical therapies may also be treatable using a 30 composition of the invention.

Compositions of the invention may also be useful to promote better or faster closure of non-healing wounds, including without limitation pressure ulcers, ulcers associated with vascular insufficiency, surgical and traumatic wounds, and the like.

Compositions of the present invention may also be involved in the generation or 35 regeneration of other tissues, such as organs (including, for example, pancreas, liver, intestine,

kidney, skin, endothelium), muscle (smooth, skeletal or cardiac) and vascular (including vascular endothelium) tissue, or for promoting the growth of cells comprising such tissues. Part of the desired effects may be by inhibition or modulation of fibrotic scarring may allow normal tissue to regenerate. A polypeptide of the present invention may also exhibit angiogenic activity.

5 A composition of the present invention may also be useful for gut protection or regeneration and treatment of lung or liver fibrosis, reperfusion injury in various tissues, and conditions resulting from systemic cytokine damage.

10 A composition of the present invention may also be useful for promoting or inhibiting differentiation of tissues described above from precursor tissues or cells; or for inhibiting the growth of tissues described above.

15 Therapeutic compositions of the invention can be used in the following:

Assays for tissue generation activity include, without limitation, those described in:

International Patent Publication No. WO95/16035 (bone, cartilage, tendon); International Patent Publication No. WO95/05846 (nerve, neuronal); International Patent Publication No.

20 WO91/07491 (skin, endothelium).

Assays for wound healing activity include, without limitation, those described in: Winter, Epidermal Wound Healing, pps. 71-112 (Maibach, H. I. and Rovee, D. T., eds.), Year Book Medical Publishers, Inc., Chicago, as modified by Eaglstein and Mertz, J. Invest. Dermatol 71:382-84 (1978).

25

#### **4.10.7 IMMUNE STIMULATING OR SUPPRESSING ACTIVITY**

A polypeptide of the present invention may also exhibit immune stimulating or immune suppressing activity, including without limitation the activities for which assays are described herein. A polynucleotide of the invention can encode a polypeptide exhibiting such activities. A 25 protein may be useful in the treatment of various immune deficiencies and disorders (including severe combined immunodeficiency (SCID)), *e.g.*, in regulating (up or down) growth and proliferation of T and/or B lymphocytes, as well as effecting the cytolytic activity of NK cells and other cell populations. These immune deficiencies may be genetic or be caused by viral (*e.g.*, HIV) as well as bacterial or fungal infections, or may result from autoimmune disorders. More 30 specifically, infectious diseases causes by viral, bacterial, fungal or other infection may be treatable using a protein of the present invention, including infections by HIV, hepatitis viruses, herpes viruses, mycobacteria, Leishmania spp., malaria spp. and various fungal infections such as candidiasis. Of course, in this regard, proteins of the present invention may also be useful where a boost to the immune system generally may be desirable, *i.e.*, in the treatment of cancer.

Autoimmune disorders which may be treated using a protein of the present invention include, for example, connective tissue disease, multiple sclerosis, systemic lupus erythematosus, rheumatoid arthritis, autoimmune pulmonary inflammation, Guillain-Barre syndrome, autoimmune thyroiditis, insulin dependent diabetes mellitus, myasthenia gravis, graft-versus-host disease and autoimmune inflammatory eye disease. Such a protein (or antagonists thereof, including antibodies) of the present invention may also be useful in the treatment of allergic reactions and conditions (e.g., anaphylaxis, serum sickness, drug reactions, food allergies, insect venom allergies, mastocytosis, allergic rhinitis, hypersensitivity pneumonitis, urticaria, angioedema, eczema, atopic dermatitis, allergic contact dermatitis, erythema multiforme, Stevens-Johnson syndrome, allergic conjunctivitis, atopic keratoconjunctivitis, venereal keratoconjunctivitis, giant papillary conjunctivitis and contact allergies), such as asthma (particularly allergic asthma) or other respiratory problems. Other conditions, in which immune suppression is desired (including, for example, organ transplantation), may also be treatable using a protein (or antagonists thereof) of the present invention. The therapeutic effects of the polypeptides or antagonists thereof on allergic reactions can be evaluated by in vivo animal models such as the cumulative contact enhancement test (Lastbom et al., Toxicology 125: 59-66, 1998), skin prick test (Hoffmann et al., Allergy 54: 446-54, 1999), guinea pig skin sensitization test (Vohr et al., Arch. Toxocol. 73: 501-9), and murine local lymph node assay (Kimber et al., J. Toxicol. Environ. Health 53: 563-79).

Using the proteins of the invention it may also be possible to modulate immune responses, in a number of ways. Down regulation may be in the form of inhibiting or blocking an immune response already in progress or may involve preventing the induction of an immune response. The functions of activated T cells may be inhibited by suppressing T cell responses or by inducing specific tolerance in T cells, or both. Immunosuppression of T cell responses is generally an active, non-antigen-specific, process which requires continuous exposure of the T cells to the suppressive agent. Tolerance, which involves inducing non-responsiveness or anergy in T cells, is distinguishable from immunosuppression in that it is generally antigen-specific and persists after exposure to the tolerizing agent has ceased. Operationally, tolerance can be demonstrated by the lack of a T cell response upon reexposure to specific antigen in the absence of the tolerizing agent.

Down regulating or preventing one or more antigen functions (including without limitation B lymphocyte antigen functions (such as, for example, B7)), e.g., preventing high level lymphokine synthesis by activated T cells, will be useful in situations of tissue, skin and organ transplantation and in graft-versus-host disease (GVHD). For example, blockage of T cell function should result in reduced tissue destruction in tissue transplantation. Typically, in tissue

transplants, rejection of the transplant is initiated through its recognition as foreign by T cells, followed by an immune reaction that destroys the transplant. The administration of a therapeutic composition of the invention may prevent cytokine synthesis by immune cells, such as T cells, and thus acts as an immunosuppressant. Moreover, a lack of costimulation may also be sufficient 5 to anergize the T cells, thereby inducing tolerance in a subject. Induction of long-term tolerance by B lymphocyte antigen-blocking reagents may avoid the necessity of repeated administration of these blocking reagents. To achieve sufficient immunosuppression or tolerance in a subject, it may also be necessary to block the function of a combination of B lymphocyte antigens.

The efficacy of particular therapeutic compositions in preventing organ transplant 10 rejection or GVHD can be assessed using animal models that are predictive of efficacy in humans. Examples of appropriate systems which can be used include allogeneic cardiac grafts in rats and xenogeneic pancreatic islet cell grafts in mice, both of which have been used to examine the immunosuppressive effects of CTLA4Ig fusion proteins *in vivo* as described in Lenschow et al., *Science* 257:789-792 (1992) and Turka et al., *Proc. Natl. Acad. Sci USA*, 89:11102-11105 15 (1992). In addition, murine models of GVHD (see Paul ed., *Fundamental Immunology*, Raven Press, New York, 1989, pp. 846-847) can be used to determine the effect of therapeutic compositions of the invention on the development of that disease.

Blocking antigen function may also be therapeutically useful for treating autoimmune 20 diseases. Many autoimmune disorders are the result of inappropriate activation of T cells that are reactive against self tissue and which promote the production of cytokines and autoantibodies involved in the pathology of the diseases. Preventing the activation of autoreactive T cells may reduce or eliminate disease symptoms. Administration of reagents which block stimulation of T cells can be used to inhibit T cell activation and prevent production of autoantibodies or T cell-derived cytokines which may be involved in the disease process. Additionally, blocking 25 reagents may induce antigen-specific tolerance of autoreactive T cells which could lead to long-term relief from the disease. The efficacy of blocking reagents in preventing or alleviating autoimmune disorders can be determined using a number of well-characterized animal models of human autoimmune diseases. Examples include murine experimental autoimmune encephalitis, systemic lupus erythmatosis in MRL/lpr/lpr mice or NZB hybrid mice, murine autoimmune 30 collagen arthritis, diabetes mellitus in NOD mice and BB rats, and murine experimental myasthenia gravis (see Paul ed., *Fundamental Immunology*, Raven Press, New York, 1989, pp. 840-856).

Upregulation of an antigen function (*e.g.*, a B lymphocyte antigen function), as a means 35 of up regulating immune responses, may also be useful in therapy. Upregulation of immune responses may be in the form of enhancing an existing immune response or eliciting an initial

immune response. For example, enhancing an immune response may be useful in cases of viral infection, including systemic viral diseases such as influenza, the common cold, and encephalitis.

Alternatively, anti-viral immune responses may be enhanced in an infected patient by  
5 removing T cells from the patient, costimulating the T cells in vitro with viral antigen-pulsed  
APCs either expressing a peptide of the present invention or together with a stimulatory form of  
a soluble peptide of the present invention and reintroducing the in vitro activated T cells into the  
patient. Another method of enhancing anti-viral immune responses would be to isolate infected  
cells from a patient, transfect them with a nucleic acid encoding a protein of the present  
10 invention as described herein such that the cells express all or a portion of the protein on their  
surface, and reintroduce the transfected cells into the patient. The infected cells would now be  
capable of delivering a costimulatory signal to, and thereby activate, T cells in vivo.

A polypeptide of the present invention may provide the necessary stimulation signal to T  
cells to induce a T cell mediated immune response against the transfected tumor cells. In  
15 addition, tumor cells which lack MHC class I or MHC class II molecules, or which fail to  
reexpress sufficient mounts of MHC class I or MHC class II molecules, can be transfected with  
nucleic acid encoding all or a portion of (e.g., a cytoplasmic-domain truncated portion) of an  
MHC class I alpha chain protein and  $\beta_2$  microglobulin protein or an MHC class II alpha chain  
protein and an MHC class II beta chain protein to thereby express MHC class I or MHC class II  
20 proteins on the cell surface. Expression of the appropriate class I or class II MHC in conjunction  
with a peptide having the activity of a B lymphocyte antigen (e.g., B7-1, B7-2, B7-3) induces a T  
cell mediated immune response against the transfected tumor cell. Optionally, a gene encoding  
an antisense construct which blocks expression of an MHC class II associated protein, such as  
the invariant chain, can also be cotransfected with a DNA encoding a peptide having the activity  
25 of a B lymphocyte antigen to promote presentation of tumor associated antigens and induce  
tumor specific immunity. Thus, the induction of a T cell mediated immune response in a human  
subject may be sufficient to overcome tumor-specific tolerance in the subject.

The activity of a protein of the invention may, among other means, be measured by the  
following methods:

30 Suitable assays for thymocyte or splenocyte cytotoxicity include, without limitation,  
those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D.  
H. Margulies, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and  
Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19;  
Chapter 7, Immunologic studies in Humans); Herrmann et al., Proc. Natl. Acad. Sci. USA  
35 78:2488-2492, 1981; Herrmann et al., J. Immunol. 128:1968-1974, 1982; Handa et al., J.

Immunol. 135:1564-1572, 1985; Takai et al., I. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988; Bowman et al., J. Virology 61:1992-1998; Bertagnolli et al., Cellular Immunology 133:327-341, 1991; Brown et al., J. Immunol. 153:3079-3092, 1994.

Assays for T-cell-dependent immunoglobulin responses and isotype switching (which 5 will identify, among others, proteins that modulate T-cell dependent antibody responses and that affect Th1/Th2 profiles) include, without limitation, those described in: Maliszewski, J. Immunol. 144:3028-3033, 1990; and Assays for B cell function: In vitro antibody production, Mond, J. J. and Brunswick, M. In *Current Protocols in Immunology*. J. E. e.a. Coligan eds. Vol 1 pp. 3.8.1-3.8.16, John Wiley and Sons, Toronto. 1994.

10 Mixed lymphocyte reaction (MLR) assays (which will identify, among others, proteins that generate predominantly Th1 and CTL responses) include, without limitation, those described in: *Current Protocols in Immunology*, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in 15 Humans); Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988; Bertagnolli et al., J. Immunol. 149:3778-3783, 1992.

20 Dendritic cell-dependent assays (which will identify, among others, proteins expressed by dendritic cells that activate naive T-cells) include, without limitation, those described in: Guery et al., J. Immunol. 134:536-544, 1995; Inaba et al., Journal of Experimental Medicine 173:549-559, 1991; Macatonia et al., Journal of Immunology 154:5071-5079, 1995; Porgador et al., Journal of Experimental Medicine 182:255-260, 1995; Nair et al., Journal of Virology 67:4062-4069, 1993; Huang et al., Science 264:961-965, 1994; Macatonia et al., Journal of Experimental Medicine 169:1255-1264, 1989; Bhardwaj et al., Journal of Clinical Investigation 94:797-807, 1994; and Inaba et al., Journal of Experimental Medicine 172:631-640, 1990.

25 Assays for lymphocyte survival/apoptosis (which will identify, among others, proteins that prevent apoptosis after superantigen induction and proteins that regulate lymphocyte homeostasis) include, without limitation, those described in: Darzynkiewicz et al., Cytometry 13:795-808, 1992; Gorczyca et al., Leukemia 7:659-670, 1993; Gorczyca et al., Cancer Research 53:1945-1951, 1993; Itoh et al., Cell 66:233-243, 1991; Zacharchuk, Journal of Immunology 30 145:4037-4045, 1990; Zamai et al., Cytometry 14:891-897, 1993; Gorczyca et al., International Journal of Oncology 1:639-648, 1992.

35 Assays for proteins that influence early steps of T-cell commitment and development include, without limitation, those described in: Antica et al., Blood 84:111-117, 1994; Fine et al., Cellular Immunology 155:111-122, 1994; Galy et al., Blood 85:2770-2778, 1995; Toki et al., Proc. Nat. Acad. Sci. USA 88:7548-7551, 1991.

#### 4.10.8 ACTIVIN/INHIBIN ACTIVITY

A polypeptide of the present invention may also exhibit activin- or inhibin-related activities. A polynucleotide of the invention may encode a polypeptide exhibiting such characteristics. Inhibins are characterized by their ability to inhibit the release of follicle stimulating hormone (FSH), while activins are characterized by their ability to stimulate the release of follicle stimulating hormone (FSH). Thus, a polypeptide of the present invention, alone or in heterodimers with a member of the inhibin family, may be useful as a contraceptive based on the ability of inhibins to decrease fertility in female mammals and decrease spermatogenesis in male mammals. Administration of sufficient amounts of other inhibins can induce infertility in these mammals. Alternatively, the polypeptide of the invention, as a homodimer or as a heterodimer with other protein subunits of the inhibin group, may be useful as a fertility inducing therapeutic, based upon the ability of activin molecules in stimulating FSH release from cells of the anterior pituitary. See, for example, U.S. Pat. No. 4,798,885. A polypeptide of the invention may also be useful for advancement of the onset of fertility in sexually immature mammals, so as to increase the lifetime reproductive performance of domestic animals such as, but not limited to, cows, sheep and pigs.

The activity of a polypeptide of the invention may, among other means, be measured by the following methods.

Assays for activin/inhibin activity include, without limitation, those described in: Vale et al., Endocrinology 91:562-572, 1972; Ling et al., Nature 321:779-782, 1986; Vale et al., Nature 321:776-779, 1986; Mason et al., Nature 318:659-663, 1985; Forage et al., Proc. Natl. Acad. Sci. USA 83:3091-3095, 1986.

#### 4.10.9 CHEMOTACTIC/CHEMOKINETIC ACTIVITY

A polypeptide of the present invention may be involved in chemotactic or chemokinetic activity for mammalian cells, including, for example, monocytes, fibroblasts, neutrophils, T-cells, mast cells, eosinophils, epithelial and/or endothelial cells. A polynucleotide of the invention can encode a polypeptide exhibiting such attributes. Chemotactic and chemokinetic receptor activation can be used to mobilize or attract a desired cell population to a desired site of action. Chemotactic or chemokinetic compositions (*e.g.* proteins, antibodies, binding partners, or modulators of the invention) provide particular advantages in treatment of wounds and other trauma to tissues, as well as in treatment of localized infections. For example, attraction of lymphocytes, monocytes or neutrophils to tumors or sites of infection may result in improved immune responses against the tumor or infecting agent.

A protein or peptide has chemotactic activity for a particular cell population if it can stimulate, directly or indirectly, the directed orientation or movement of such cell population. Preferably, the protein or peptide has the ability to directly stimulate directed movement of cells. Whether a particular protein has chemotactic activity for a population of cells can be readily 5 determined by employing such protein or peptide in any known assay for cell chemotaxis.

Therapeutic compositions of the invention can be used in the following:

Assays for chemotactic activity (which will identify proteins that induce or prevent chemotaxis) consist of assays that measure the ability of a protein to induce the migration of cells across a membrane as well as the ability of a protein to induce the adhesion of one cell 10 population to another cell population. Suitable assays for movement and adhesion include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Marguiles, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 6.12, Measurement of alpha and beta Chemokines 6.12.1-6.12.28; Taub et al. J. Clin. Invest. 95:1370-1376, 1995; Lind et al. APMIS 103:140-146, 15 1995; Muller et al Eur. J. Immunol. 25:1744-1748; Gruber et al. J. of Immunol. 152:5860-5867, 1994; Johnston et al. J. of Immunol. 153:1762-1768, 1994.

#### **4.10.10 HEMOSTATIC AND THROMBOLYTIC ACTIVITY**

A polypeptide of the invention may also be involved in hemostasis or thrombolysis or 20 thrombosis. A polynucleotide of the invention can encode a polypeptide exhibiting such attributes. Compositions may be useful in treatment of various coagulation disorders (including hereditary disorders, such as hemophilias) or to enhance coagulation and other hemostatic events in treating wounds resulting from trauma, surgery or other causes. A composition of the invention may also be useful for dissolving or inhibiting formation of thromboses and for 25 treatment and prevention of conditions resulting therefrom (such as, for example, infarction of cardiac and central nervous system vessels (e.g., stroke).

Therapeutic compositions of the invention can be used in the following:

Assay for hemostatic and thrombolytic activity include, without limitation, those 30 described in: Linet et al., J. Clin. Pharmacol. 26:131-140, 1986; Burdick et al., Thrombosis Res. 45:413-419, 1987; Humphrey et al., Fibrinolysis 5:71-79 (1991); Schaub, Prostaglandins 35:467-474, 1988.

#### **4.10.11 CANCER DIAGNOSIS AND THERAPY**

Polypeptides of the invention may be involved in cancer cell generation, proliferation or 35 metastasis. Detection of the presence or amount of polynucleotides or polypeptides of the

invention may be useful for the diagnosis and/or prognosis of one or more types of cancer. For example, the presence or increased expression of a polynucleotide/polypeptide of the invention may indicate a hereditary risk of cancer, a precancerous condition, or an ongoing malignancy. Conversely, a defect in the gene or absence of the polypeptide may be associated with a cancer 5 condition. Identification of single nucleotide polymorphisms associated with cancer or a predisposition to cancer may also be useful for diagnosis or prognosis.

Cancer treatments promote tumor regression by inhibiting tumor cell proliferation, inhibiting angiogenesis (growth of new blood vessels that is necessary to support tumor growth) and/or prohibiting metastasis by reducing tumor cell motility or invasiveness. Therapeutic 10 compositions of the invention may be effective in adult and pediatric oncology including in solid phase tumors/malignancies, locally advanced tumors, human soft tissue sarcomas, metastatic cancer, including lymphatic metastases, blood cell malignancies including multiple myeloma, acute and chronic leukemias, and lymphomas, head and neck cancers including mouth cancer, larynx cancer and thyroid cancer, lung cancers including small cell carcinoma and non-small cell 15 cancers, breast cancers including small cell carcinoma and ductal carcinoma, gastrointestinal cancers including esophageal cancer, stomach cancer, colon cancer, colorectal cancer and polyps associated with colorectal neoplasia, pancreatic cancers, liver cancer, urologic cancers including bladder cancer and prostate cancer, malignancies of the female genital tract including ovarian carcinoma, uterine (including endometrial) cancers, and solid tumor in the ovarian follicle, 20 kidney cancers including renal cell carcinoma, brain cancers including intrinsic brain tumors, neuroblastoma, astrocytic brain tumors, gliomas, metastatic tumor cell invasion in the central nervous system, bone cancers including osteomas, skin cancers including malignant melanoma, tumor progression of human skin keratinocytes, squamous cell carcinoma, basal cell carcinoma, hemangiopericytoma and Karposi's sarcoma.

25 Polypeptides, polynucleotides, or modulators of polypeptides of the invention (including inhibitors and stimulators of the biological activity of the polypeptide of the invention) may be administered to treat cancer. Therapeutic compositions can be administered in therapeutically effective dosages alone or in combination with adjuvant cancer therapy such as surgery, chemotherapy, radiotherapy, thermotherapy, and laser therapy, and may provide a beneficial 30 effect, *e.g.* reducing tumor size, slowing rate of tumor growth, inhibiting metastasis, or otherwise improving overall clinical condition, without necessarily eradicating the cancer.

The composition can also be administered in therapeutically effective amounts as a portion of an anti-cancer cocktail. An anti-cancer cocktail is a mixture of the polypeptide or modulator of the invention with one or more anti-cancer drugs in addition to a pharmaceutically 35 acceptable carrier for delivery. The use of anti-cancer cocktails as a cancer treatment is routine.

Anti-cancer drugs that are well known in the art and can be used as a treatment in combination with the polypeptide or modulator of the invention include: Actinomycin D, Aminoglutethimide, Asparaginase, Bleomycin, Busulfan, Carboplatin, Carmustine, Chlorambucil, Cisplatin (cis-DDP), Cyclophosphamide, Cytarabine HCl (Cytosine arabinoside), Dacarbazine, Dactinomycin, 5 Daunorubicin HCl, Doxorubicin HCl, Estramustine phosphate sodium, Etoposide (V16-213), Floxuridine, 5-Fluorouracil (5-Fu), Flutamide, Hydroxyurea (hydroxycarbamide), Ifosfamide, Interferon Alpha-2a, Interferon Alpha-2b, Leuprolide acetate (LHRH-releasing factor analog), Lomustine, Mechlorethamine HCl (nitrogen mustard), Melphalan, Mercaptopurine, Mesna, Methotrexate (MTX), Mitomycin, Mitoxantrone HCl, Octreotide, Plicamycin, Procarbazine HCl, 10 Streptozocin, Tamoxifen citrate, Thioguanine, Thiotepa, Vinblastine sulfate, Vincristine sulfate, Amsacrine, Azacitidine, Hexamethylmelamine, Interleukin-2, Mitoguazone, Pentostatin, Semustine, Teniposide, and Vindesine sulfate.

In addition, therapeutic compositions of the invention may be used for prophylactic treatment of cancer. There are hereditary conditions and/or environmental situations (e.g. 15 exposure to carcinogens) known in the art that predispose an individual to developing cancers. Under these circumstances, it may be beneficial to treat these individuals with therapeutically effective doses of the polypeptide of the invention to reduce the risk of developing cancers.

*In vitro* models can be used to determine the effective doses of the polypeptide of the invention as a potential cancer treatment. These *in vitro* models include proliferation assays of 20 cultured tumor cells, growth of cultured tumor cells in soft agar (see Freshney, (1987) Culture of Animal Cells: A Manual of Basic Technique, Wiley-Liss, New York, NY Ch 18 and Ch 21), tumor systems in nude mice as described in Giovanella et al., J. Natl. Can. Inst., 52: 921-30 (1974), mobility and invasive potential of tumor cells in Boyden Chamber assays as described in Pilkington et al., Anticancer Res., 17: 4107-9 (1997), and angiogenesis assays such as induction 25 of vascularization of the chick chorioallantoic membrane or induction of vascular endothelial cell migration as described in Ribatta et al., Intl. J. Dev. Biol., 40: 1189-97 (1999) and Li et al., Clin. Exp. Metastasis, 17:423-9 (1999), respectively. Suitable tumor cells lines are available, e.g. from American Type Tissue Culture Collection catalogs.

#### 30 4.10.12 RECEPTOR/LIGAND ACTIVITY

A polypeptide of the present invention may also demonstrate activity as receptor, receptor ligand or inhibitor or agonist of receptor/ligand interactions. A polynucleotide of the invention can encode a polypeptide exhibiting such characteristics. Examples of such receptors and ligands include, without limitation, cytokine receptors and their ligands, receptor kinases and 35 their ligands, receptor phosphatases and their ligands, receptors involved in cell-cell interactions

and their ligands (including without limitation, cellular adhesion molecules (such as selectins, integrins and their ligands) and receptor/ligand pairs involved in antigen presentation, antigen recognition and development of cellular and humoral immune responses. Receptors and ligands are also useful for screening of potential peptide or small molecule inhibitors of the relevant 5 receptor/ligand interaction. A protein of the present invention (including, without limitation, fragments of receptors and ligands) may themselves be useful as inhibitors of receptor/ligand interactions.

The activity of a polypeptide of the invention may, among other means, be measured by the following methods:

10 Suitable assays for receptor-ligand activity include without limitation those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley- Interscience (Chapter 7.28, Measurement of Cellular Adhesion under static conditions 7.28.1- 7.28.22), Takai et al., Proc. Natl. Acad. Sci. USA 84:6864-6868, 1987; Bierer et al., J. Exp. Med. 168:1145-1156, 1988; 15 Rosenstein et al., J. Exp. Med. 169:149-160 1989; Stoltenborg et al., J. Immunol. Methods 175:59-68, 1994; Stitt et al., Cell 80:661-670, 1995.

20 By way of example, the polypeptides of the invention may be used as a receptor for a ligand(s) thereby transmitting the biological activity of that ligand(s). Ligands may be identified through binding assays, affinity chromatography, dihybrid screening assays, BIACore assays, gel overlay assays, or other methods known in the art.

Studies characterizing drugs or proteins as agonist or antagonist or partial agonists or a partial antagonist require the use of other proteins as competing ligands. The polypeptides of the present invention or ligand(s) thereof may be labeled by being coupled to radioisotopes, colorimetric molecules or a toxin molecules by conventional methods. ("Guide to Protein 25 Purification" Murray P. Deutscher (ed) Methods in Enzymology Vol. 182 (1990) Academic Press, Inc. San Diego). Examples of radioisotopes include, but are not limited to, tritium and carbon-14. Examples of colorimetric molecules include, but are not limited to, fluorescent molecules such as fluorescamine, or rhodamine or other colorimetric molecules. Examples of toxins include, but are not limited, to ricin.

30

#### 4.10.13 DRUG SCREENING

This invention is particularly useful for screening chemical compounds by using the novel polypeptides or binding fragments thereof in any of a variety of drug screening techniques. The polypeptides or fragments employed in such a test may either be free in solution, affixed to a 35 solid support, borne on a cell surface or located intracellularly. One method of drug screening

utilizes eukaryotic or prokaryotic host cells which are stably transformed with recombinant nucleic acids expressing the polypeptide or a fragment thereof. Drugs are screened against such transformed cells in competitive binding assays. Such cells, either in viable or fixed form, can be used for standard binding assays. One may measure, for example, the formation of 5 complexes between polypeptides of the invention or fragments and the agent being tested or examine the diminution in complex formation between the novel polypeptides and an appropriate cell line, which are well known in the art.

Sources for test compounds that may be screened for ability to bind to or modulate (*i.e.*, increase or decrease) the activity of polypeptides of the invention include (1) inorganic and 10 organic chemical libraries, (2) natural product libraries, and (3) combinatorial libraries comprised of either random or mimetic peptides, oligonucleotides or organic molecules.

Chemical libraries may be readily synthesized or purchased from a number of commercial sources, and may include structural analogs of known compounds or compounds that are identified as "hits" or "leads" via natural product screening.

15 The sources of natural product libraries are microorganisms (including bacteria and fungi), animals, plants or other vegetation, or marine organisms, and libraries of mixtures for screening may be created by: (1) fermentation and extraction of broths from soil, plant or marine microorganisms or (2) extraction of the organisms themselves. Natural product libraries include polyketides, non-ribosomal peptides, and (non-naturally occurring) variants thereof. For a 20 review, see *Science* 282:63-68 (1998).

Combinatorial libraries are composed of large numbers of peptides, oligonucleotides or 25 organic compounds and can be readily prepared by traditional automated synthesis methods, PCR, cloning or proprietary synthetic methods. Of particular interest are peptide and oligonucleotide combinatorial libraries. Still other libraries of interest include peptide, protein, peptidomimetic, multiparallel synthetic collection, recombinatorial, and polypeptide libraries. For a review of combinatorial chemistry and libraries created therefrom, see Myers, *Curr. Opin. Biotechnol.* 8:701-707 (1997). For reviews and examples of peptidomimetic libraries, see Al-Obeidi et al., *Mol. Biotechnol.*, 9(3):205-23 (1998); Hruby et al., *Curr Opin Chem Biol.* 1(1):114-19 (1997); Dorner et al., *Bioorg Med Chem*, 4(5):709-15 (1996) (alkylated dipeptides).

30 Identification of modulators through use of the various libraries described herein permits modification of the candidate "hit" (or "lead") to optimize the capacity of the "hit" to bind a polypeptide of the invention. The molecules identified in the binding assay are then tested for antagonist or agonist activity in *in vivo* tissue culture or animal models that are well known in the art. In brief, the molecules are titrated into a plurality of cell cultures or animals and then tested 35 for either cell/animal death or prolonged survival of the animal/cells.

The binding molecules thus identified may be complexed with toxins, *e.g.*, ricin or cholera, or with other compounds that are toxic to cells such as radioisotopes. The toxin-binding molecule complex is then targeted to a tumor or other cell by the specificity of the binding molecule for a polypeptide of the invention. Alternatively, the binding molecules may be 5 complexed with imaging agents for targeting and imaging purposes.

#### 4.10.14 ASSAY FOR RECEPTOR ACTIVITY

The invention also provides methods to detect specific binding of a polypeptide *e.g.* a ligand or a receptor. The art provides numerous assays particularly useful for identifying 10 previously unknown binding partners for receptor polypeptides of the invention. For example, expression cloning using mammalian or bacterial cells, or dihybrid screening assays can be used to identify polynucleotides encoding binding partners. As another example, affinity chromatography with the appropriate immobilized polypeptide of the invention can be used to isolate polypeptides that recognize and bind polypeptides of the invention. There are a number 15 of different libraries used for the identification of compounds, and in particular small molecules, that modulate (*i.e.*, increase or decrease) biological activity of a polypeptide of the invention. Ligands for receptor polypeptides of the invention can also be identified by adding exogenous 20 ligands, or cocktails of ligands to two cell populations that are genetically identical except for the expression of the receptor of the invention: one cell population expresses the receptor of the invention whereas the other does not. The response of the two cell populations to the addition of ligands(s) are then compared. Alternatively, an expression library can be co-expressed with the 25 polypeptide of the invention in cells and assayed for an autocrine response to identify potential ligand(s). As still another example, BIACore assays, gel overlay assays, or other methods known in the art can be used to identify binding partner polypeptides, including, (1) organic and inorganic chemical libraries, (2) natural product libraries, and (3) combinatorial libraries comprised of random peptides, oligonucleotides or organic molecules.

The role of downstream intracellular signaling molecules in the signaling cascade of the polypeptide of the invention can be determined. For example, a chimeric protein in which the cytoplasmic domain of the polypeptide of the invention is fused to the extracellular portion of a 30 protein, whose ligand has been identified, is produced in a host cell. The cell is then incubated with the ligand specific for the extracellular portion of the chimeric protein, thereby activating the chimeric receptor. Known downstream proteins involved in intracellular signaling can then be assayed for expected modifications *i.e.* phosphorylation. Other methods known to those in the art can also be used to identify signaling molecules involved in receptor activity.

**4.10.15 ANTI-INFLAMMATORY ACTIVITY**

Compositions of the present invention may also exhibit anti-inflammatory activity. The anti-inflammatory activity may be achieved by providing a stimulus to cells involved in the inflammatory response, by inhibiting or promoting cell-cell interactions (such as, for example, cell adhesion), by inhibiting or promoting chemotaxis of cells involved in the inflammatory process, inhibiting or promoting cell extravasation, or by stimulating or suppressing production of other factors which more directly inhibit or promote an inflammatory response. Compositions with such activities can be used to treat inflammatory conditions including chronic or acute conditions), including without limitation intimation associated with infection (such as septic shock, sepsis or systemic inflammatory response syndrome (SIRS)), ischemia-reperfusion injury, endotoxin lethality, arthritis, complement-mediated hyperacute rejection, nephritis, cytokine or chemokine-induced lung injury, inflammatory bowel disease, Crohn's disease or resulting from over production of cytokines such as TNF or IL-1. Compositions of the invention may also be useful to treat anaphylaxis and hypersensitivity to an antigenic substance or material.

Compositions of this invention may be utilized to prevent or treat conditions such as, but not limited to, sepsis, acute pancreatitis, endotoxin shock, cytokine induced shock, rheumatoid arthritis, chronic inflammatory arthritis, pancreatic cell damage from diabetes mellitus type 1, graft versus host disease, inflammatory bowel disease, inflammation associated with pulmonary disease, other autoimmune disease or inflammatory disease, an antiproliferative agent such as for acute or chronic myelogenous leukemia or in the prevention of premature labor secondary to intrauterine infections.

**4.10.16 LEUKEMIAS**

Leukemias and related disorders may be treated or prevented by administration of a therapeutic that promotes or inhibits function of the polynucleotides and/or polypeptides of the invention. Such leukemias and related disorders include but are not limited to acute leukemia, acute lymphocytic leukemia, acute myelocytic leukemia, myeloblastic, promyelocytic, myelomonocytic, monocytic, erythroleukemia, chronic leukemia, chronic myelocytic (granulocytic) leukemia and chronic lymphocytic leukemia (for a review of such disorders, see Fishman et al., 1985, Medicine, 2d Ed., J.B. Lippincott Co., Philadelphia).

**4.10.17 NERVOUS SYSTEM DISORDERS**

Nervous system disorders, involving cell types which can be tested for efficacy of intervention with compounds that modulate the activity of the polynucleotides and/or polypeptides of the invention, and which can be treated upon thus observing an indication of

therapeutic utility, include but are not limited to nervous system injuries, and diseases or disorders which result in either a disconnection of axons, a diminution or degeneration of neurons, or demyelination. Nervous system lesions which may be treated in a patient (including human and non-human mammalian patients) according to the invention include but are not limited to the following lesions of either the central (including spinal cord, brain) or peripheral nervous systems:

- 5 (i) traumatic lesions, including lesions caused by physical injury or associated with surgery, for example, lesions which sever a portion of the nervous system, or compression injuries;
- 10 (ii) ischemic lesions, in which a lack of oxygen in a portion of the nervous system results in neuronal injury or death, including cerebral infarction or ischemia, or spinal cord infarction or ischemia;
- 15 (iii) infectious lesions, in which a portion of the nervous system is destroyed or injured as a result of infection, for example, by an abscess or associated with infection by human immunodeficiency virus, herpes zoster, or herpes simplex virus or with Lyme disease, tuberculosis, syphilis;
- 20 (iv) degenerative lesions, in which a portion of the nervous system is destroyed or injured as a result of a degenerative process including but not limited to degeneration associated with Parkinson's disease, Alzheimer's disease, Huntington's chorea, or amyotrophic lateral sclerosis;
- 25 (v) lesions associated with nutritional diseases or disorders, in which a portion of the nervous system is destroyed or injured by a nutritional disorder or disorder of metabolism including but not limited to, vitamin B12 deficiency, folic acid deficiency, Wernicke disease, tobacco-alcohol amblyopia, Marchiafava-Bignami disease (primary degeneration of the corpus callosum), and alcoholic cerebellar degeneration;
- 30 (vi) neurological lesions associated with systemic diseases including but not limited to diabetes (diabetic neuropathy, Bell's palsy), systemic lupus erythematosus, carcinoma, or sarcoidosis;
- (vii) lesions caused by toxic substances including alcohol, lead, or particular neurotoxins; and
- (viii) demyelinated lesions in which a portion of the nervous system is destroyed or injured by a demyelinating disease including but not limited to multiple sclerosis, human immunodeficiency virus-associated myelopathy, transverse myelopathy or various etiologies, progressive multifocal leukoencephalopathy, and central pontine myelinolysis.

Therapeutics which are useful according to the invention for treatment of a nervous system disorder may be selected by testing for biological activity in promoting the survival or differentiation of neurons. For example, and not by way of limitation, therapeutics which elicit any of the following effects may be useful according to the invention:

5 (i) increased survival time of neurons in culture;  
(ii) increased sprouting of neurons in culture or *in vivo*;  
(iii) increased production of a neuron-associated molecule in culture or *in vivo*, *e.g.*, choline acetyltransferase or acetylcholinesterase with respect to motor neurons; or  
(iv) decreased symptoms of neuron dysfunction *in vivo*.

10 Such effects may be measured by any method known in the art. In preferred, non-limiting embodiments, increased survival of neurons may be measured by the method set forth in Arakawa et al. (1990, *J. Neurosci.* 10:3507-3515); increased sprouting of neurons may be detected by methods set forth in Pestronk et al. (1980, *Exp. Neurol.* 70:65-82) or Brown et al. (1981, *Ann. Rev. Neurosci.* 4:17-42); increased production of neuron-associated molecules may 15 be measured by bioassay, enzymatic assay, antibody binding, Northern blot assay, *etc.*, depending on the molecule to be measured; and motor neuron dysfunction may be measured by assessing the physical manifestation of motor neuron disorder, *e.g.*, weakness, motor neuron conduction velocity, or functional disability.

20 In specific embodiments, motor neuron disorders that may be treated according to the invention include but are not limited to disorders such as infarction, infection, exposure to toxin, trauma, surgical damage, degenerative disease or malignancy that may affect motor neurons as well as other components of the nervous system, as well as disorders that selectively affect neurons such as amyotrophic lateral sclerosis, and including but not limited to progressive spinal muscular atrophy, progressive bulbar palsy, primary lateral sclerosis, infantile and juvenile 25 muscular atrophy, progressive bulbar paralysis of childhood (Fazio-Londe syndrome), poliomyelitis and the post polio syndrome, and Hereditary Motorsensory Neuropathy (Charcot-Marie-Tooth Disease).

#### 4.10.18 OTHER ACTIVITIES

30 A polypeptide of the invention may also exhibit one or more of the following additional activities or effects: inhibiting the growth, infection or function of, or killing, infectious agents, including, without limitation, bacteria, viruses, fungi and other parasites; effecting (suppressing or enhancing) bodily characteristics, including, without limitation, height, weight, hair color, eye color, skin, fat to lean ratio or other tissue pigmentation, or organ or body part size or shape 35 (such as, for example, breast augmentation or diminution, change in bone form or shape);

effecting biorhythms or circadian cycles or rhythms; effecting the fertility of male or female subjects; effecting the metabolism, catabolism, anabolism, processing, utilization, storage or elimination of dietary fat, lipid, protein, carbohydrate, vitamins, minerals, co-factors or other nutritional factors or component(s); effecting behavioral characteristics, including, without limitation, appetite, libido, stress, cognition (including cognitive disorders), depression (including depressive disorders) and violent behaviors; providing analgesic effects or other pain reducing effects; promoting differentiation and growth of embryonic stem cells in lineages other than hematopoietic lineages; hormonal or endocrine activity; in the case of enzymes, correcting deficiencies of the enzyme and treating deficiency-related diseases; treatment of hyperproliferative disorders (such as, for example, psoriasis); immunoglobulin-like activity (such as, for example, the ability to bind antigens or complement); and the ability to act as an antigen in a vaccine composition to raise an immune response against such protein or another material or entity which is cross-reactive with such protein.

15 **4.10.19 IDENTIFICATION OF POLYMORPHISMS**

The demonstration of polymorphisms makes possible the identification of such polymorphisms in human subjects and the pharmacogenetic use of this information for diagnosis and treatment. Such polymorphisms may be associated with, *e.g.*, differential predisposition or susceptibility to various disease states (such as disorders involving inflammation or immune response) or a differential response to drug administration, and this genetic information can be used to tailor preventive or therapeutic treatment appropriately. For example, the existence of a polymorphism associated with a predisposition to inflammation or autoimmune disease makes possible the diagnosis of this condition in humans by identifying the presence of the polymorphism.

25 Polymorphisms can be identified in a variety of ways known in the art which all generally involve obtaining a sample from a patient, analyzing DNA from the sample, optionally involving isolation or amplification of the DNA, and identifying the presence of the polymorphism in the DNA. For example, PCR may be used to amplify an appropriate fragment of genomic DNA which may then be sequenced. Alternatively, the DNA may be subjected to 30 allele-specific oligonucleotide hybridization (in which appropriate oligonucleotides are hybridized to the DNA under conditions permitting detection of a single base mismatch) or to a single nucleotide extension assay (in which an oligonucleotide that hybridizes immediately adjacent to the position of the polymorphism is extended with one or more labeled nucleotides). In addition, traditional restriction fragment length polymorphism analysis (using restriction 35 enzymes that provide differential digestion of the genomic DNA depending on the presence or

absence of the polymorphism) may be performed. Arrays with nucleotide sequences of the present invention can be used to detect polymorphisms. The array can comprise modified nucleotide sequences of the present invention in order to detect the nucleotide sequences of the present invention. In the alternative, any one of the nucleotide sequences of the present invention can be placed on the array to detect changes from those sequences.

5 Alternatively a polymorphism resulting in a change in the amino acid sequence could also be detected by detecting a corresponding change in amino acid sequence of the protein, *e.g.*, by an antibody specific to the variant sequence.

10 **4.10.20 ARTHRITIS AND INFLAMMATION**

The immunosuppressive effects of the compositions of the invention against rheumatoid arthritis is determined in an experimental animal model system. The experimental model system is adjuvant induced arthritis in rats, and the protocol is described by J. Holoskitz, et al., 1983, *Science*, 219:56, or by B. Waksman et al., 1963, *Int. Arch. Allergy Appl. Immunol.*, 23:129.

15 Induction of the disease can be caused by a single injection, generally intradermally, of a suspension of killed *Mycobacterium tuberculosis* in complete Freund's adjuvant (CFA). The route of injection can vary, but rats may be injected at the base of the tail with an adjuvant mixture. The polypeptide is administered in phosphate buffered solution (PBS) at a dose of about 1-5 mg/kg. The control consists of administering PBS only.

20 The procedure for testing the effects of the test compound would consist of intradermally injecting killed *Mycobacterium tuberculosis* in CFA followed by immediately administering the test compound and subsequent treatment every other day until day 24. At 14, 15, 18, 20, 22, and 24 days after injection of *Mycobacterium* CFA, an overall arthritis score may be obtained as described by J. Holoskitz above. An analysis of the data would reveal that the test compound

25 would have a dramatic affect on the swelling of the joints as measured by a decrease of the arthritis score.

**4.11 THERAPEUTIC METHODS**

The compositions (including polypeptide fragments, analogs, variants and antibodies or other binding partners or modulators including antisense polynucleotides) of the invention have numerous applications in a variety of therapeutic methods. Examples of therapeutic applications include, but are not limited to, those exemplified herein.

**4.11.1 EXAMPLE**

One embodiment of the invention is the administration of an effective amount of the polypeptides or other composition of the invention to individuals affected by a disease or disorder that can be modulated by regulating the peptides of the invention. While the mode of administration is not particularly important, parenteral administration is preferred. An exemplary mode of administration is to deliver an intravenous bolus. The dosage of the polypeptides or other composition of the invention will normally be determined by the prescribing physician. It is to be expected that the dosage will vary according to the age, weight, condition and response of the individual patient. Typically, the amount of polypeptide administered per dose will be in the range of about 0.01 $\mu$ g/kg to 100 mg/kg of body weight, with the preferred dose being about 0.1 $\mu$ g/kg to 10 mg/kg of patient body weight. For parenteral administration, polypeptides of the invention will be formulated in an injectable form combined with a pharmaceutically acceptable parenteral vehicle. Such vehicles are well known in the art and examples include water, saline, Ringer's solution, dextrose solution, and solutions consisting of small amounts of the human serum albumin. The vehicle may contain minor amounts of additives that maintain the isotonicity and stability of the polypeptide or other active ingredient. The preparation of such solutions is within the skill of the art.

#### 4.12 PHARMACEUTICAL FORMULATIONS AND ROUTES OF ADMINISTRATION

A protein or other composition of the present invention (from whatever source derived, including without limitation from recombinant and non-recombinant sources and including antibodies and other binding partners of the polypeptides of the invention) may be administered to a patient in need, by itself, or in pharmaceutical compositions where it is mixed with suitable carriers or excipient(s) at doses to treat or ameliorate a variety of disorders. Such a composition may optionally contain (in addition to protein or other active ingredient and a carrier) diluents, fillers, salts, buffers, stabilizers, solubilizers, and other materials well known in the art. The term "pharmaceutically acceptable" means a non-toxic material that does not interfere with the effectiveness of the biological activity of the active ingredient(s). The characteristics of the carrier will depend on the route of administration. The pharmaceutical composition of the invention may also contain cytokines, lymphokines, or other hematopoietic factors such as M-CSF, GM-CSF, TNF, IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-11, IL-12, IL-13, IL-14, IL-15, IFN, TNF0, TNF1, TNF2, G-CSF, Meg-CSF, thrombopoietin, stem cell factor, and erythropoietin. In further compositions, proteins of the invention may be combined with other agents beneficial to the treatment of the disease or disorder in question. These agents include various growth factors such as epidermal growth factor (EGF), platelet-derived growth

factor (PDGF), transforming growth factors (TGF- $\alpha$  and TGF- $\beta$ ), insulin-like growth factor (IGF), as well as cytokines described herein.

The pharmaceutical composition may further contain other agents which either enhance the activity of the protein or other active ingredient or complement its activity or use in treatment. Such additional factors and/or agents may be included in the pharmaceutical composition to produce a synergistic effect with protein or other active ingredient of the invention, or to minimize side effects. Conversely, protein or other active ingredient of the present invention may be included in formulations of the particular clotting factor, cytokine, lymphokine, other hematopoietic factor, thrombolytic or anti-thrombotic factor, or anti-inflammatory agent to minimize side effects of the clotting factor, cytokine, lymphokine, other hematopoietic factor, thrombolytic or anti-thrombotic factor, or anti-inflammatory agent (such as IL-1Ra, IL-1 Hy1, IL-1 Hy2, anti-TNF, corticosteroids, immunosuppressive agents). A protein of the present invention may be active in multimers (*e.g.*, heterodimers or homodimers) or complexes with itself or other proteins. As a result, pharmaceutical compositions of the invention may comprise a protein of the invention in such multimeric or complexed form.

As an alternative to being included in a pharmaceutical composition of the invention including a first protein, a second protein or a therapeutic agent may be concurrently administered with the first protein (*e.g.*, at the same time, or at differing times provided that therapeutic concentrations of the combination of agents is achieved at the treatment site).

Techniques for formulation and administration of the compounds of the instant application may be found in "Remington's Pharmaceutical Sciences," Mack Publishing Co., Easton, PA, latest edition. A therapeutically effective dose further refers to that amount of the compound sufficient to result in amelioration of symptoms, *e.g.*, treatment, healing, prevention or amelioration of the relevant medical condition, or an increase in rate of treatment, healing, prevention or amelioration of such conditions. When applied to an individual active ingredient, administered alone, a therapeutically effective dose refers to that ingredient alone. When applied to a combination, a therapeutically effective dose refers to combined amounts of the active ingredients that result in the therapeutic effect, whether administered in combination, serially or simultaneously.

In practicing the method of treatment or use of the present invention, a therapeutically effective amount of protein or other active ingredient of the present invention is administered to a mammal having a condition to be treated. Protein or other active ingredient of the present invention may be administered in accordance with the method of the invention either alone or in combination with other therapies such as treatments employing cytokines, lymphokines or other hematopoietic factors. When co-administered with one or more cytokines, lymphokines or other

hematopoietic factors, protein or other active ingredient of the present invention may be administered either simultaneously with the cytokine(s), lymphokine(s), other hematopoietic factor(s), thrombolytic or anti-thrombotic factors, or sequentially. If administered sequentially, the attending physician will decide on the appropriate sequence of administering protein or other active ingredient of the present invention in combination with cytokine(s), lymphokine(s), other hematopoietic factor(s), thrombolytic or anti-thrombotic factors.

#### **4.12.1 ROUTES OF ADMINISTRATION**

Suitable routes of administration may, for example, include oral, rectal, transmucosal, or 10 intestinal administration; parenteral delivery, including intramuscular, subcutaneous, intramedullary injections, as well as intrathecal, direct intraventricular, intravenous, intraperitoneal, intranasal, or intraocular injections. Administration of protein or other active ingredient of the present invention used in the pharmaceutical composition or to practice the method of the present invention can be carried out in a variety of conventional ways, such as oral 15 ingestion, inhalation, topical application or cutaneous, subcutaneous, intraperitoneal, parenteral or intravenous injection. Intravenous administration to the patient is preferred.

Alternately, one may administer the compound in a local rather than systemic manner, for example, via injection of the compound directly into a arthritic joints or in fibrotic tissue, often in a depot or sustained release formulation. In order to prevent the scarring process frequently 20 occurring as complication of glaucoma surgery, the compounds may be administered topically, for example, as eye drops. Furthermore, one may administer the drug in a targeted drug delivery system, for example, in a liposome coated with a specific antibody, targeting, for example, arthritic or fibrotic tissue. The liposomes will be targeted to and taken up selectively by the afflicted tissue.

25 The polypeptides of the invention are administered by any route that delivers an effective dosage to the desired site of action. The determination of a suitable route of administration and an effective dosage for a particular indication is within the level of skill in the art. Preferably for wound treatment, one administers the therapeutic compound directly to the site. Suitable dosage ranges for the polypeptides of the invention can be extrapolated from these dosages or from 30 similar studies in appropriate animal models. Dosages can then be adjusted as necessary by the clinician to provide maximal therapeutic benefit.

#### **4.12.2 COMPOSITIONS/FORMULATIONS**

35 Pharmaceutical compositions for use in accordance with the present invention thus may be formulated in a conventional manner using one or more physiologically acceptable carriers

comprising excipients and auxiliaries which facilitate processing of the active compounds into preparations which can be used pharmaceutically. These pharmaceutical compositions may be manufactured in a manner that is itself known, *e.g.*, by means of conventional mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping or 5 lyophilizing processes. Proper formulation is dependent upon the route of administration chosen. When a therapeutically effective amount of protein or other active ingredient of the present invention is administered orally, protein or other active ingredient of the present invention will be in the form of a tablet, capsule, powder, solution or elixir. When administered in tablet form, the pharmaceutical composition of the invention may additionally contain a solid 10 carrier such as a gelatin or an adjuvant. The tablet, capsule, and powder contain from about 5 to 95% protein or other active ingredient of the present invention, and preferably from about 25 to 90% protein or other active ingredient of the present invention. When administered in liquid form, a liquid carrier such as water, petroleum, oils of animal or plant origin such as peanut oil, mineral oil, soybean oil, or sesame oil, or synthetic oils may be added. The liquid form of the 15 pharmaceutical composition may further contain physiological saline solution, dextrose or other saccharide solution, or glycols such as ethylene glycol, propylene glycol or polyethylene glycol. When administered in liquid form, the pharmaceutical composition contains from about 0.5 to 90% by weight of protein or other active ingredient of the present invention, and preferably from about 1 to 50% protein or other active ingredient of the present invention.

20 When a therapeutically effective amount of protein or other active ingredient of the present invention is administered by intravenous, cutaneous or subcutaneous injection, protein or other active ingredient of the present invention will be in the form of a pyrogen-free, parenterally acceptable aqueous solution. The preparation of such parenterally acceptable protein or other active ingredient solutions, having due regard to pH, isotonicity, stability, and the like, is within 25 the skill in the art. A preferred pharmaceutical composition for intravenous, cutaneous, or subcutaneous injection should contain, in addition to protein or other active ingredient of the present invention, an isotonic vehicle such as Sodium Chloride Injection, Ringer's Injection, Dextrose Injection, Dextrose and Sodium Chloride Injection, Lactated Ringer's Injection, or other vehicle as known in the art. The pharmaceutical composition of the present invention may 30 also contain stabilizers, preservatives, buffers, antioxidants, or other additives known to those of skill in the art. For injection, the agents of the invention may be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks's solution, Ringer's solution, or physiological saline buffer. For transmucosal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known 35 in the art.

For oral administration, the compounds can be formulated readily by combining the active compounds with pharmaceutically acceptable carriers well known in the art. Such carriers enable the compounds of the invention to be formulated as tablets, pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions and the like, for oral ingestion by a patient to be 5 treated. Pharmaceutical preparations for oral use can be obtained from a solid excipient, optionally grinding a resulting mixture, and processing the mixture of granules, after adding suitable auxiliaries, if desired, to obtain tablets or dragee cores. Suitable excipients are, in particular, fillers such as sugars, including lactose, sucrose, mannitol, or sorbitol; cellulose preparations such as, for example, maize starch, wheat starch, rice starch, potato starch, gelatin, 10 gum tragacanth, methyl cellulose, hydroxypropylmethyl-cellulose, sodium carboxymethylcellulose, and/or polyvinylpyrrolidone (PVP). If desired, disintegrating agents may be added, such as the cross-linked polyvinyl pyrrolidone, agar, or alginic acid or a salt thereof such as sodium alginate. Dragee cores are provided with suitable coatings. For this purpose, concentrated sugar solutions may be used, which may optionally contain gum arabic, 15 talc, polyvinyl pyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures. Dyestuffs or pigments may be added to the tablets or dragee coatings for identification or to characterize different combinations of active compound doses.

Pharmaceutical preparations which can be used orally include push-fit capsules made of 20 gelatin, as well as soft, sealed capsules made of gelatin and a plasticizer, such as glycerol or sorbitol. The push-fit capsules can contain the active ingredients in admixture with filler such as lactose, binders such as starches, and/or lubricants such as talc or magnesium stearate and, optionally, stabilizers. In soft capsules, the active compounds may be dissolved or suspended in suitable liquids, such as fatty oils, liquid paraffin, or liquid polyethylene glycols. In addition, 25 stabilizers may be added. All formulations for oral administration should be in dosages suitable for such administration. For buccal administration, the compositions may take the form of tablets or lozenges formulated in conventional manner.

For administration by inhalation, the compounds for use according to the present invention are conveniently delivered in the form of an aerosol spray presentation from 30 pressurized packs or a nebuliser, with the use of a suitable propellant, *e.g.*, dichlorodifluoromethane, trichlorofluoromethane, dichlorotetrafluoroethane, carbon dioxide or other suitable gas. In the case of a pressurized aerosol the dosage unit may be determined by providing a valve to deliver a metered amount. Capsules and cartridges of, *e.g.*, gelatin for use in an inhaler or insufflator may be formulated containing a powder mix of the compound and a 35 suitable powder base such as lactose or starch. The compounds may be formulated for parenteral

administration by injection, *e.g.*, by bolus injection or continuous infusion. Formulations for injection may be presented in unit dosage form, *e.g.*, in ampules or in multi-dose containers, with an added preservative. The compositions may take such forms as suspensions, solutions or emulsions in oily or aqueous vehicles, and may contain formulatory agents such as suspending, 5 stabilizing and/or dispersing agents.

Pharmaceutical formulations for parenteral administration include aqueous solutions of the active compounds in water-soluble form. Additionally, suspensions of the active compounds may be prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate or 10 triglycerides, or liposomes. Aqueous injection suspensions may contain substances which increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Optionally, the suspension may also contain suitable stabilizers or agents which increase the solubility of the compounds to allow for the preparation of highly concentrated 15 solutions. Alternatively, the active ingredient may be in powder form for constitution with a suitable vehicle, *e.g.*, sterile pyrogen-free water, before use.

The compounds may also be formulated in rectal compositions such as suppositories or retention enemas, *e.g.*, containing conventional suppository bases such as cocoa butter or other glycerides. In addition to the formulations described previously, the compounds may also be formulated as a depot preparation. Such long acting formulations may be administered by 20 implantation (for example subcutaneously or intramuscularly) or by intramuscular injection. Thus, for example, the compounds may be formulated with suitable polymeric or hydrophobic materials (for example as an emulsion in an acceptable oil) or ion exchange resins, or as sparingly soluble derivatives, for example, as a sparingly soluble salt.

A pharmaceutical carrier for the hydrophobic compounds of the invention is a co-solvent 25 system comprising benzyl alcohol, a nonpolar surfactant, a water-miscible organic polymer, and an aqueous phase. The co-solvent system may be the VPD co-solvent system. VPD is a solution of 3% w/v benzyl alcohol, 8% w/v of the nonpolar surfactant polysorbate 80, and 65% w/v polyethylene glycol 300, made up to volume in absolute ethanol. The VPD co-solvent system (VPD:5W) consists of VPD diluted 1:1 with a 5% dextrose in water solution. This co-solvent 30 system dissolves hydrophobic compounds well, and itself produces low toxicity upon systemic administration. Naturally, the proportions of a co-solvent system may be varied considerably without destroying its solubility and toxicity characteristics. Furthermore, the identity of the co-solvent components may be varied: for example, other low-toxicity nonpolar surfactants may be used instead of polysorbate 80; the fraction size of polyethylene glycol may be varied; other 35 biocompatible polymers may replace polyethylene glycol, *e.g.* polyvinyl pyrrolidone; and other

sugars or polysaccharides may substitute for dextrose. Alternatively, other delivery systems for hydrophobic pharmaceutical compounds may be employed. Liposomes and emulsions are well known examples of delivery vehicles or carriers for hydrophobic drugs. Certain organic solvents such as dimethylsulfoxide also may be employed, although usually at the cost of greater toxicity.

5    Additionally, the compounds may be delivered using a sustained-release system, such as semipermeable matrices of solid hydrophobic polymers containing the therapeutic agent. Various types of sustained-release materials have been established and are well known by those skilled in the art. Sustained-release capsules may, depending on their chemical nature, release the compounds for a few weeks up to over 100 days. Depending on the chemical nature and the  
10    biological stability of the therapeutic reagent, additional strategies for protein or other active ingredient stabilization may be employed.

The pharmaceutical compositions also may comprise suitable solid or gel phase carriers or excipients. Examples of such carriers or excipients include but are not limited to calcium carbonate, calcium phosphate, various sugars, starches, cellulose derivatives, gelatin, and  
15    polymers such as polyethylene glycols. Many of the active ingredients of the invention may be provided as salts with pharmaceutically compatible counter ions. Such pharmaceutically acceptable base addition salts are those salts which retain the biological effectiveness and properties of the free acids and which are obtained by reaction with inorganic or organic bases such as sodium hydroxide, magnesium hydroxide, ammonia, trialkylamine, dialkylamine,  
20    monoalkylamine, dibasic amino acids, sodium acetate, potassium benzoate, triethanol amine and the like.

The pharmaceutical composition of the invention may be in the form of a complex of the protein(s) or other active ingredient(s) of present invention along with protein or peptide antigens. The protein and/or peptide antigen will deliver a stimulatory signal to both B and T  
25    lymphocytes. B lymphocytes will respond to antigen through their surface immunoglobulin receptor. T lymphocytes will respond to antigen through the T cell receptor (TCR) following presentation of the antigen by MHC proteins. MHC and structurally related proteins including those encoded by class I and class II MHC genes on host cells will serve to present the peptide antigen(s) to T lymphocytes. The antigen components could also be supplied as purified  
30    MHC-peptide complexes alone or with co-stimulatory molecules that can directly signal T cells. Alternatively antibodies able to bind surface immunoglobulin and other molecules on B cells as well as antibodies able to bind the TCR and other molecules on T cells can be combined with the pharmaceutical composition of the invention.

The pharmaceutical composition of the invention may be in the form of a liposome in  
35    which protein of the present invention is combined, in addition to other pharmaceutically

acceptable carriers, with amphipathic agents such as lipids which exist in aggregated form as micelles, insoluble monolayers, liquid crystals, or lamellar layers in aqueous solution. Suitable lipids for liposomal formulation include, without limitation, monoglycerides, diglycerides, sulfatides, lysolecithins, phospholipids, saponin, bile acids, and the like. Preparation of such 5 liposomal formulations is within the level of skill in the art, as disclosed, for example, in U.S. Patent Nos. 4,235,871; 4,501,728; 4,837,028; and 4,737,323, all of which are incorporated herein by reference.

The amount of protein or other active ingredient of the present invention in the pharmaceutical composition of the present invention will depend upon the nature and severity of 10 the condition being treated, and on the nature of prior treatments which the patient has undergone. Ultimately, the attending physician will decide the amount of protein or other active ingredient of the present invention with which to treat each individual patient. Initially, the attending physician will administer low doses of protein or other active ingredient of the present invention and observe the patient's response. Larger doses of protein or other active ingredient 15 of the present invention may be administered until the optimal therapeutic effect is obtained for the patient, and at that point the dosage is not increased further. It is contemplated that the various pharmaceutical compositions used to practice the method of the present invention should contain about 0.01  $\mu$ g to about 100 mg (preferably about 0.1  $\mu$ g to about 10 mg, more preferably about 0.1  $\mu$ g to about 1 mg) of protein or other active ingredient of the present invention per kg 20 body weight. For compositions of the present invention which are useful for bone, cartilage, tendon or ligament regeneration, the therapeutic method includes administering the composition topically, systematically, or locally as an implant or device. When administered, the therapeutic composition for use in this invention is, of course, in a pyrogen-free, physiologically acceptable form. Further, the composition may desirably be encapsulated or injected in a viscous form for 25 delivery to the site of bone, cartilage or tissue damage. Topical administration may be suitable for wound healing and tissue repair. Therapeutically useful agents other than a protein or other active ingredient of the invention which may also optionally be included in the composition as described above, may alternatively or additionally, be administered simultaneously or sequentially with the composition in the methods of the invention. Preferably for bone and/or 30 cartilage formation, the composition would include a matrix capable of delivering the protein-containing or other active ingredient-containing composition to the site of bone and/or cartilage damage, providing a structure for the developing bone and cartilage and optimally capable of being resorbed into the body. Such matrices may be formed of materials presently in use for other implanted medical applications.

The choice of matrix material is based on biocompatibility, biodegradability, mechanical properties, cosmetic appearance and interface properties. The particular application of the compositions will define the appropriate formulation. Potential matrices for the compositions may be biodegradable and chemically defined calcium sulfate, tricalcium phosphate, 5 hydroxyapatite, polylactic acid, polyglycolic acid and polyanhydrides. Other potential materials are biodegradable and biologically well-defined, such as bone or dermal collagen. Further matrices are comprised of pure proteins or extracellular matrix components. Other potential matrices are nonbiodegradable and chemically defined, such as sintered hydroxyapatite, bioglass, aluminates, or other ceramics. Matrices may be comprised of combinations of any of the above 10 mentioned types of material, such as polylactic acid and hydroxyapatite or collagen and tricalcium phosphate. The bioceramics may be altered in composition, such as in calcium-aluminate-phosphate and processing to alter pore size, particle size, particle shape, and biodegradability. Presently preferred is a 50:50 (mole weight) copolymer of lactic acid and glycolic acid in the form of porous particles having diameters ranging from 150 to 800 microns. 15 In some applications, it will be useful to utilize a sequestering agent, such as carboxymethyl cellulose or autologous blood clot, to prevent the protein compositions from disassociating from the matrix.

A preferred family of sequestering agents is cellulosic materials such as alkylcelluloses (including hydroxyalkylcelluloses), including methylcellulose, ethylcellulose, 20 hydroxyethylcellulose, hydroxypropylcellulose, hydroxypropyl-methylcellulose, and carboxymethylcellulose, the most preferred being cationic salts of carboxymethylcellulose (CMC). Other preferred sequestering agents include hyaluronic acid, sodium alginate, poly(ethylene glycol), polyoxyethylene oxide, carboxyvinyl polymer and poly(vinyl alcohol). The amount of sequestering agent useful herein is 0.5-20 wt %, preferably 1-10 wt % based on 25 total formulation weight, which represents the amount necessary to prevent desorption of the protein from the polymer matrix and to provide appropriate handling of the composition, yet not so much that the progenitor cells are prevented from infiltrating the matrix, thereby providing the protein the opportunity to assist the osteogenic activity of the progenitor cells. In further compositions, proteins or other active ingredients of the invention may be combined with other 30 agents beneficial to the treatment of the bone and/or cartilage defect, wound, or tissue in question. These agents include various growth factors such as epidermal growth factor (EGF), platelet derived growth factor (PDGF), transforming growth factors (TGF- $\alpha$  and TGF- $\beta$ ), and insulin-like growth factor (IGF).

The therapeutic compositions are also presently valuable for veterinary applications. 35 Particularly domestic animals and thoroughbred horses, in addition to humans, are desired

patients for such treatment with proteins or other active ingredients of the present invention. The dosage regimen of a protein-containing pharmaceutical composition to be used in tissue regeneration will be determined by the attending physician considering various factors which modify the action of the proteins, *e.g.*, amount of tissue weight desired to be formed, the site of 5 damage, the condition of the damaged tissue, the size of a wound, type of damaged tissue (*e.g.*, bone), the patient's age, sex, and diet, the severity of any infection, time of administration and other clinical factors. The dosage may vary with the type of matrix used in the reconstitution and with inclusion of other proteins in the pharmaceutical composition. For example, the 10 addition of other known growth factors, such as IGF I (insulin like growth factor I), to the final composition, may also effect the dosage. Progress can be monitored by periodic assessment of tissue/bone growth and/or repair, for example, X-rays, histomorphometric determinations and tetracycline labeling.

20 Polynucleotides of the present invention can also be used for gene therapy. Such polynucleotides can be introduced either *in vivo* or *ex vivo* into cells for expression in a 15 mammalian subject. Polynucleotides of the invention may also be administered by other known methods for introduction of nucleic acid into a cell or organism (including, without limitation, in the form of viral vectors or naked DNA). Cells may also be cultured *ex vivo* in the presence of proteins of the present invention in order to proliferate or to produce a desired effect on or activity in such cells. Treated cells can then be introduced *in vivo* for therapeutic purposes.

20

#### **4.12.3 EFFECTIVE DOSAGE**

Pharmaceutical compositions suitable for use in the present invention include 25 compositions wherein the active ingredients are contained in an effective amount to achieve its intended purpose. More specifically, a therapeutically effective amount means an amount effective to prevent development of or to alleviate the existing symptoms of the subject being treated. Determination of the effective amount is well within the capability of those skilled in the art, especially in light of the detailed disclosure provided herein. For any compound used in the method of the invention, the therapeutically effective dose can be estimated initially from appropriate *in vitro* assays. For example, a dose can be formulated in animal models to achieve a 30 circulating concentration range that can be used to more accurately determine useful doses in humans. For example, a dose can be formulated in animal models to achieve a circulating concentration range that includes the IC<sub>50</sub> as determined in cell culture (*i.e.*, the concentration of the test compound which achieves a half-maximal inhibition of the protein's biological activity). Such information can be used to more accurately determine useful doses in humans.

A therapeutically effective dose refers to that amount of the compound that results in amelioration of symptoms or a prolongation of survival in a patient. Toxicity and therapeutic efficacy of such compounds can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, *e.g.*, for determining the LD<sub>50</sub> (the dose lethal to 50% of the population) and the ED<sub>50</sub> (the dose therapeutically effective in 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index and it can be expressed as the ratio between LD<sub>50</sub> and ED<sub>50</sub>. Compounds which exhibit high therapeutic indices are preferred. The data obtained from these cell culture assays and animal studies can be used in formulating a range of dosage for use in human. The dosage of such compounds lies preferably within a range of circulating concentrations that include the ED<sub>50</sub> with little or no toxicity. The dosage may vary within this range depending upon the dosage form employed and the route of administration utilized. The exact formulation, route of administration and dosage can be chosen by the individual physician in view of the patient's condition. See, *e.g.*, Fingl et al., 1975, in "The Pharmacological Basis of Therapeutics", Ch. 1 p.1. Dosage amount and interval may be adjusted individually to provide plasma levels of the active moiety which are sufficient to maintain the desired effects, or minimal effective concentration (MEC). The MEC will vary for each compound but can be estimated from *in vitro* data. Dosages necessary to achieve the MEC will depend on individual characteristics and route of administration. However, HPLC assays or bioassays can be used to determine plasma concentrations.

Dosage intervals can also be determined using MEC value. Compounds should be administered using a regimen which maintains plasma levels above the MEC for 10-90% of the time, preferably between 30-90% and most preferably between 50-90%. In cases of local administration or selective uptake, the effective local concentration of the drug may not be related to plasma concentration.

An exemplary dosage regimen for polypeptides or other compositions of the invention will be in the range of about 0.01  $\mu$ g/kg to 100 mg/kg of body weight daily, with the preferred dose being about 0.1  $\mu$ g/kg to 25 mg/kg of patient body weight daily, varying in adults and children. Dosing may be once daily, or equivalent doses may be delivered at longer or shorter intervals.

The amount of composition administered will, of course, be dependent on the subject being treated, on the subject's age and weight, the severity of the affliction, the manner of administration and the judgment of the prescribing physician.

#### 4.12.4 PACKAGING

The compositions may, if desired, be presented in a pack or dispenser device which may contain one or more unit dosage forms containing the active ingredient. The pack may, for example, comprise metal or plastic foil, such as a blister pack. The pack or dispenser device may be accompanied by instructions for administration. Compositions comprising a compound of the 5 invention formulated in a compatible pharmaceutical carrier may also be prepared, placed in an appropriate container, and labeled for treatment of an indicated condition.

#### 4.13 ANTIBODIES

Also included in the invention are antibodies to proteins, or fragments of proteins of the 10 invention. The term "antibody" as used herein refers to immunoglobulin molecules and immunologically active portions of immunoglobulin (Ig) molecules, *i.e.*, molecules that contain an antigen binding site that specifically binds (immunoreacts with) an antigen. Such antibodies include, but are not limited to, polyclonal, monoclonal, chimeric, single chain,  $F_{ab}$ ,  $F_{ab}'$  and  $F_{(ab')2}$  fragments, and an  $F_{ab}$  expression library. In general, an antibody molecule obtained from 15 humans relates to any of the classes IgG, IgM, IgA, IgE and IgD, which differ from one another by the nature of the heavy chain present in the molecule. Certain classes have subclasses as well, such as IgG<sub>1</sub>, IgG<sub>2</sub>, and others. Furthermore, in humans, the light chain may be a kappa chain or a lambda chain. Reference herein to antibodies includes a reference to all such classes, subclasses and types of human antibody species.

20 An isolated related protein of the invention may be intended to serve as an antigen, or a portion or fragment thereof, and additionally can be used as an immunogen to generate antibodies that immunospecifically bind the antigen, using standard techniques for polyclonal and monoclonal antibody preparation. The full-length protein can be used or, alternatively, the invention provides antigenic peptide fragments of the antigen for use as immunogens. An 25 antigenic peptide fragment comprises at least 6 amino acid residues of the amino acid sequence of the full length protein, (for example the amino acid sequence shown in SEQ ID NO: 1010), and encompasses an epitope thereof such that an antibody raised against the peptide forms a specific immune complex with the full length protein or with any fragment that contains the epitope. Preferably, the antigenic peptide comprises at least 10 amino acid residues, or at least 30 30 amino acid residues, or at least 20 amino acid residues, or at least 30 amino acid residues. Preferred epitopes encompassed by the antigenic peptide are regions of the protein that are located on its surface; commonly these are hydrophilic regions.

In certain embodiments of the invention, at least one epitope encompassed by the antigenic peptide is a region of -related protein that is located on the surface of the protein, *e.g.*, a 35 hydrophilic region. A hydrophobicity analysis of the human related protein sequence will

indicate which regions of a related protein are particularly hydrophilic and, therefore, are likely to encode surface residues useful for targeting antibody production. As a means for targeting antibody production, hydropathy plots showing regions of hydrophilicity and hydrophobicity may be generated by any method well known in the art, including, for example, the Kyte 5 Doolittle or the Hopp Woods methods, either with or without Fourier transformation. See, e.g., Hopp and Woods, 1981, *Proc. Nat. Acad. Sci. USA* 78: 3824-3828; Kyte and Doolittle 1982, *J. Mol. Biol.* 157: 105-142, each of which is incorporated herein by reference in its entirety. Antibodies that are specific for one or more domains within an antigenic protein, or derivatives, fragments, analogs or homologs thereof, are also provided herein.

10 A protein of the invention, or a derivative, fragment, analog, homolog or ortholog thereof, may be utilized as an immunogen in the generation of antibodies that immunospecifically bind these protein components.

Various procedures known within the art may be used for the production of polyclonal or monoclonal antibodies directed against a protein of the invention, or against derivatives, 15 fragments, analogs homologs or orthologs thereof (see, for example, *Antibodies: A Laboratory Manual*, Harlow E, and Lane D, 1988, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, incorporated herein by reference). Some of these antibodies are discussed below.

### 5.13.1 Polyclonal Antibodies

20 For the production of polyclonal antibodies, various suitable host animals (e.g., rabbit, goat, mouse or other mammal) may be immunized by one or more injections with the native protein, a synthetic variant thereof, or a derivative of the foregoing. An appropriate immunogenic preparation can contain, for example, the naturally occurring immunogenic protein, a chemically synthesized polypeptide representing the immunogenic protein, or a 25 recombinantly expressed immunogenic protein. Furthermore, the protein may be conjugated to a second protein known to be immunogenic in the mammal being immunized. Examples of such immunogenic proteins include but are not limited to keyhole limpet hemocyanin, serum albumin, bovine thyroglobulin, and soybean trypsin inhibitor. The preparation can further include an adjuvant. Various adjuvants used to increase the immunological response include, but are not 30 limited to, Freund's (complete and incomplete), mineral gels (e.g., aluminum hydroxide), surface active substances (e.g., lyssolecithin, pluronic polyols, polyanions, peptides, oil emulsions, dinitrophenol, etc.), adjuvants usable in humans such as Bacille Calmette-Guerin and *Corynebacterium parvum*, or similar immunostimulatory agents. Additional examples of adjuvants which can be employed include MPL-TDM adjuvant (monophosphoryl Lipid A, 35 synthetic trehalose dicorynomycolate).

The polyclonal antibody molecules directed against the immunogenic protein can be isolated from the mammal (e.g., from the blood) and further purified by well known techniques, such as affinity chromatography using protein A or protein G, which provide primarily the IgG fraction of immune serum. Subsequently, or alternatively, the specific antigen which is the 5 target of the immunoglobulin sought, or an epitope thereof, may be immobilized on a column to purify the immune specific antibody by immunoaffinity chromatography. Purification of immunoglobulins is discussed, for example, by D. Wilkinson (The Scientist, published by The Scientist, Inc., Philadelphia PA, Vol. 14, No. 8 (April 17, 2000), pp. 25-28).

10 **5.13.2 Monoclonal Antibodies**

The term "monoclonal antibody" (MAb) or "monoclonal antibody composition", as used herein, refers to a population of antibody molecules that contain only one molecular species of antibody molecule consisting of a unique light chain gene product and a unique heavy chain gene product. In particular, the complementarity determining regions (CDRs) of the monoclonal 15 antibody are identical in all the molecules of the population. MAbs thus contain an antigen binding site capable of immunoreacting with a particular epitope of the antigen characterized by a unique binding affinity for it.

Monoclonal antibodies can be prepared using hybridoma methods, such as those described by Kohler and Milstein, Nature, 256:495 (1975). In a hybridoma method, a mouse, 20 hamster, or other appropriate host animal, is typically immunized with an immunizing agent to elicit lymphocytes that produce or are capable of producing antibodies that will specifically bind to the immunizing agent. Alternatively, the lymphocytes can be immunized in vitro.

The immunizing agent will typically include the protein antigen, a fragment thereof or a fusion protein thereof. Generally, either peripheral blood lymphocytes are used if cells of human 25 origin are desired, or spleen cells or lymph node cells are used if non-human mammalian sources are desired. The lymphocytes are then fused with an immortalized cell line using a suitable fusing agent, such as polyethylene glycol, to form a hybridoma cell (Goding, Monoclonal Antibodies: Principles and Practice, Academic Press, (1986) pp. 59-103). Immortalized cell lines are usually transformed mammalian cells, particularly myeloma cells of rodent, bovine and 30 human origin. Usually, rat or mouse myeloma cell lines are employed. The hybridoma cells can be cultured in a suitable culture medium that preferably contains one or more substances that inhibit the growth or survival of the unfused, immortalized cells. For example, if the parental cells lack the enzyme hypoxanthine guanine phosphoribosyl transferase (HGPRT or HPRT), the culture medium for the hybridomas typically will include hypoxanthine, aminopterin, and 35 thymidine ("HAT medium"), which substances prevent the growth of HGPRT-deficient cells.

Preferred immortalized cell lines are those that fuse efficiently, support stable high level expression of antibody by the selected antibody-producing cells, and are sensitive to a medium such as HAT medium. More preferred immortalized cell lines are murine myeloma lines, which can be obtained, for instance, from the Salk Institute Cell Distribution Center, San Diego, 5 California and the American Type Culture Collection, Manassas, Virginia. Human myeloma and mouse-human heteromyeloma cell lines also have been described for the production of human monoclonal antibodies (Kozbor, J. Immunol., 133:3001 (1984); Brodeur et al., Monoclonal Antibody Production Techniques and Applications, Marcel Dekker, Inc., New York, (1987) pp. 51-63).

10 The culture medium in which the hybridoma cells are cultured can then be assayed for the presence of monoclonal antibodies directed against the antigen. Preferably, the binding specificity of monoclonal antibodies produced by the hybridoma cells is determined by immunoprecipitation or by an in vitro binding assay, such as radioimmunoassay (RIA) or enzyme-linked immunoabsorbent assay (ELISA). Such techniques and assays are known in the 15 art. The binding affinity of the monoclonal antibody can, for example, be determined by the Scatchard analysis of Munson and Pollard, Anal. Biochem., 107:220 (1980). Preferably, antibodies having a high degree of specificity and a high binding affinity for the target antigen are isolated.

20 After the desired hybridoma cells are identified, the clones can be subcloned by limiting dilution procedures and grown by standard methods. Suitable culture media for this purpose include, for example, Dulbecco's Modified Eagle's Medium and RPMI-1640 medium. Alternatively, the hybridoma cells can be grown in vivo as ascites in a mammal.

25 The monoclonal antibodies secreted by the subclones can be isolated or purified from the culture medium or ascites fluid by conventional immunoglobulin purification procedures such as, for example, protein A-Sepharose, hydroxylapatite chromatography, gel electrophoresis, dialysis, or affinity chromatography.

30 The monoclonal antibodies can also be made by recombinant DNA methods, such as those described in U.S. Patent No. 4,816,567. DNA encoding the monoclonal antibodies of the invention can be readily isolated and sequenced using conventional procedures (e.g., by using oligonucleotide probes that are capable of binding specifically to genes encoding the heavy and light chains of murine antibodies). The hybridoma cells of the invention serve as a preferred 35 source of such DNA. Once isolated, the DNA can be placed into expression vectors, which are then transfected into host cells such as simian COS cells, Chinese hamster ovary (CHO) cells, or myeloma cells that do not otherwise produce immunoglobulin protein, to obtain the synthesis of monoclonal antibodies in the recombinant host cells. The DNA also can be modified, for

example, by substituting the coding sequence for human heavy and light chain constant domains in place of the homologous murine sequences (U.S. Patent No. 4,816,567; Morrison, Nature **368**, 812-13 (1994)) or by covalently joining to the immunoglobulin coding sequence all or part of the coding sequence for a non-immunoglobulin polypeptide. Such a non-immunoglobulin 5 polypeptide can be substituted for the constant domains of an antibody of the invention, or can be substituted for the variable domains of one antigen-combining site of an antibody of the invention to create a chimeric bivalent antibody.

### **5.13.2 Humanized Antibodies**

10 The antibodies directed against the protein antigens of the invention can further comprise humanized antibodies or human antibodies. These antibodies are suitable for administration to humans without engendering an immune response by the human against the administered immunoglobulin. Humanized forms of antibodies are chimeric immunoglobulins, immunoglobulin chains or fragments thereof (such as Fv, Fab, Fab', F(ab')<sub>2</sub> or other antigen-binding subsequences of antibodies) that are principally comprised of the sequence of a human immunoglobulin, and contain minimal sequence derived from a non-human immunoglobulin. Humanization can be performed following the method of Winter and co-workers (Jones et al., Nature, **321**:522-525 (1986); Riechmann et al., Nature, **332**:323-327 (1988); Verhoeyen et al., Science, **239**:1534-1536 (1988)), by substituting rodent CDRs or CDR sequences for the 15 corresponding sequences of a human antibody. (See also U.S. Patent No. 5,225,539.) In some instances, Fv framework residues of the human immunoglobulin are replaced by corresponding non-human residues. Humanized antibodies can also comprise residues which are found neither in the recipient antibody nor in the imported CDR or framework sequences. In general, the humanized antibody will comprise substantially all of at least one, and typically two, variable 20 domains, in which all or substantially all of the CDR regions correspond to those of a non-human immunoglobulin and all or substantially all of the framework regions are those of a human immunoglobulin consensus sequence. The humanized antibody optimally also will comprise at 25 least a portion of an immunoglobulin constant region (Fc), typically that of a human immunoglobulin (Jones et al., 1986; Riechmann et al., 1988; and Presta, Curr. Op. Struct. Biol., **2**:593-596 (1992)).

### **5.13.3 Human Antibodies**

30 Fully human antibodies relate to antibody molecules in which essentially the entire sequences of both the light chain and the heavy chain, including the CDRs, arise from human genes. Such antibodies are termed "human antibodies", or "fully human antibodies" herein.

Human monoclonal antibodies can be prepared by the trioma technique; the human B-cell hybridoma technique (see Kozbor, et al., 1983 *Immunol Today* 4: 72) and the EBV hybridoma technique to produce human monoclonal antibodies (see Cole, et al., 1985 In: *MONOCLONAL ANTIBODIES AND CANCER THERAPY*, Alan R. Liss, Inc., pp. 77-96). Human monoclonal antibodies may be utilized in the practice of the present invention and may be produced by using human hybridomas (see Cote, et al., 1983. *Proc Natl Acad Sci USA* 80: 2026-2030) or by transforming human B-cells with Epstein Barr Virus in vitro (see Cole, et al., 1985 In: *MONOCLONAL ANTIBODIES AND CANCER THERAPY*, Alan R. Liss, Inc., pp. 77-96).

In addition, human antibodies can also be produced using additional techniques, 10 including phage display libraries (Hoogenboom and Winter, *J. Mol. Biol.*, 227:381 (1991); Marks et al., *J. Mol. Biol.*, 222:581 (1991)). Similarly, human antibodies can be made by introducing human immunoglobulin loci into transgenic animals, e.g., mice in which the endogenous immunoglobulin genes have been partially or completely inactivated. Upon challenge, human antibody production is observed, which closely resembles that seen in humans 15 in all respects, including gene rearrangement, assembly, and antibody repertoire. This approach is described, for example, in U.S. Patent Nos. 5,545,807; 5,545,806; 5,569,825; 5,625,126; 5,633,425; 5,661,016, and in Marks et al. (*Bio/Technology* 10, 779-783 (1992)); Lonberg et al. (*Nature* 368 856-859 (1994)); Morrison (*Nature* 368, 812-13 (1994)); Fishwild et al., (*Nature Biotechnology* 14, 845-51 (1996)); Neuberger (*Nature Biotechnology* 14, 826 (1996)); and 20 Lonberg and Huszar (*Intern. Rev. Immunol.* 13 65-93 (1995)).

Human antibodies may additionally be produced using transgenic nonhuman animals which are modified so as to produce fully human antibodies rather than the animal's endogenous antibodies in response to challenge by an antigen. (See PCT publication WO94/02602). The endogenous genes encoding the heavy and light immunoglobulin chains in the nonhuman host 25 have been incapacitated, and active loci encoding human heavy and light chain immunoglobulins are inserted into the host's genome. The human genes are incorporated, for example, using yeast artificial chromosomes containing the requisite human DNA segments. An animal which provides all the desired modifications is then obtained as progeny by crossbreeding intermediate transgenic animals containing fewer than the full complement of the modifications. The 30 preferred embodiment of such a nonhuman animal is a mouse, and is termed the Xenomouse<sup>TM</sup> as disclosed in PCT publications WO 96/33735 and WO 96/34096. This animal produces B cells which secrete fully human immunoglobulins. The antibodies can be obtained directly from the animal after immunization with an immunogen of interest, as, for example, a preparation of a polyclonal antibody, or alternatively from immortalized B cells derived from the animal, such as 35 hybridomas producing monoclonal antibodies. Additionally, the genes encoding the

immunoglobulins with human variable regions can be recovered and expressed to obtain the antibodies directly, or can be further modified to obtain analogs of antibodies such as, for example, single chain Fv molecules.

An example of a method of producing a nonhuman host, exemplified as a mouse, lacking expression of an endogenous immunoglobulin heavy chain is disclosed in U.S. Patent No. 5,939,598. It can be obtained by a method including deleting the J segment genes from at least one endogenous heavy chain locus in an embryonic stem cell to prevent rearrangement of the locus and to prevent formation of a transcript of a rearranged immunoglobulin heavy chain locus, the deletion being effected by a targeting vector containing a gene encoding a selectable marker; and producing from the embryonic stem cell a transgenic mouse whose somatic and germ cells contain the gene encoding the selectable marker.

A method for producing an antibody of interest, such as a human antibody, is disclosed in U.S. Patent No. 5,916,771. It includes introducing an expression vector that contains a nucleotide sequence encoding a heavy chain into one mammalian host cell in culture, introducing an expression vector containing a nucleotide sequence encoding a light chain into another mammalian host cell, and fusing the two cells to form a hybrid cell. The hybrid cell expresses an antibody containing the heavy chain and the light chain.

In a further improvement on this procedure, a method for identifying a clinically relevant epitope on an immunogen, and a correlative method for selecting an antibody that binds immunospecifically to the relevant epitope with high affinity, are disclosed in PCT publication WO 99/53049.

#### 5.13.4 F<sub>ab</sub> Fragments and Single Chain Antibodies

According to the invention, techniques can be adapted for the production of single-chain antibodies specific to an antigenic protein of the invention (see e.g., U.S. Patent No. 4,946,778). In addition, methods can be adapted for the construction of F<sub>ab</sub> expression libraries (see e.g., Huse, et al., 1989 Science 246: 1275-1281) to allow rapid and effective identification of monoclonal F<sub>ab</sub> fragments with the desired specificity for a protein or derivatives, fragments, analogs or homologs thereof. Antibody fragments that contain the idiotypes to a protein antigen may be produced by techniques known in the art including, but not limited to: (i) an F<sub>(ab')2</sub> fragment produced by pepsin digestion of an antibody molecule; (ii) an F<sub>ab</sub> fragment generated by reducing the disulfide bridges of an F<sub>(ab')2</sub> fragment; (iii) an F<sub>ab</sub> fragment generated by the treatment of the antibody molecule with papain and a reducing agent and (iv) F<sub>v</sub> fragments.

Bispecific antibodies are monoclonal, preferably human or humanized, antibodies that have binding specificities for at least two different antigens. In the present case, one of the binding specificities is for an antigenic protein of the invention. The second binding target is any other antigen, and advantageously is a cell-surface protein or receptor or receptor subunit.

5 Methods for making bispecific antibodies are known in the art. Traditionally, the recombinant production of bispecific antibodies is based on the co-expression of two immunoglobulin heavy-chain/light-chain pairs, where the two heavy chains have different specificities (Milstein and Cuello, Nature, 305:537-539 (1983)). Because of the random assortment of immunoglobulin heavy and light chains, these hybridomas (quadromas) produce a  
10 potential mixture of ten different antibody molecules, of which only one has the correct bispecific structure. The purification of the correct molecule is usually accomplished by affinity chromatography steps. Similar procedures are disclosed in WO 93/08829, published 13 May 1993, and in Traunecker *et al.*, 1991 EMBO J., 10:3655-3659.

15 Antibody variable domains with the desired binding specificities (antibody-antigen combining sites) can be fused to immunoglobulin constant domain sequences. The fusion preferably is with an immunoglobulin heavy-chain constant domain, comprising at least part of the hinge, CH2, and CH3 regions. It is preferred to have the first heavy-chain constant region (CH1) containing the site necessary for light-chain binding present in at least one of the fusions. DNAs encoding the immunoglobulin heavy-chain fusions and, if desired, the immunoglobulin  
20 light chain, are inserted into separate expression vectors, and are co-transfected into a suitable host organism. For further details of generating bispecific antibodies see, for example, Suresh *et al.*, Methods in Enzymology, 121:210 (1986).

25 According to another approach described in WO 96/27011, the interface between a pair of antibody molecules can be engineered to maximize the percentage of heterodimers which are recovered from recombinant cell culture. The preferred interface comprises at least a part of the CH3 region of an antibody constant domain. In this method, one or more small amino acid side chains from the interface of the first antibody molecule are replaced with larger side chains (e.g. tyrosine or tryptophan). Compensatory "cavities" of identical or similar size to the large side chain(s) are created on the interface of the second antibody molecule by replacing large amino  
30 acid side chains with smaller ones (e.g. alanine or threonine). This provides a mechanism for increasing the yield of the heterodimer over other unwanted end-products such as homodimers.

35 Bispecific antibodies can be prepared as full length antibodies or antibody fragments (e.g. F(ab')<sub>2</sub> bispecific antibodies). Techniques for generating bispecific antibodies from antibody fragments have been described in the literature. For example, bispecific antibodies can be prepared using chemical linkage. Brennan *et al.*, Science 229:81 (1985) describe a procedure

wherein intact antibodies are proteolytically cleaved to generate  $F(ab')_2$  fragments. These fragments are reduced in the presence of the dithiol complexing agent sodium arsenite to stabilize vicinal dithiols and prevent intermolecular disulfide formation. The  $Fab'$  fragments generated are then converted to thionitrobenzoate (TNB) derivatives. One of the  $Fab'$ -TNB derivatives is then reconverted to the  $Fab'$ -thiol by reduction with mercaptoethylamine and is mixed with an equimolar amount of the other  $Fab'$ -TNB derivative to form the bispecific antibody. The bispecific antibodies produced can be used as agents for the selective immobilization of enzymes.

Additionally,  $Fab'$  fragments can be directly recovered from *E. coli* and chemically coupled to form bispecific antibodies. Shalaby et al., *J. Exp. Med.* 175:217-225 (1992) describe the production of a fully humanized bispecific antibody  $F(ab')_2$  molecule. Each  $Fab'$  fragment was separately secreted from *E. coli* and subjected to directed chemical coupling in vitro to form the bispecific antibody. The bispecific antibody thus formed was able to bind to cells overexpressing the ErbB2 receptor and normal human T cells, as well as trigger the lytic activity of human cytotoxic lymphocytes against human breast tumor targets.

Various techniques for making and isolating bispecific antibody fragments directly from recombinant cell culture have also been described. For example, bispecific antibodies have been produced using leucine zippers. Kostelny et al., *J. Immunol.* 148(5):1547-1553 (1992). The leucine zipper peptides from the Fos and Jun proteins were linked to the  $Fab'$  portions of two different antibodies by gene fusion. The antibody homodimers were reduced at the hinge region to form monomers and then re-oxidized to form the antibody heterodimers. This method can also be utilized for the production of antibody homodimers. The "diabody" technology described by Hollinger et al., *Proc. Natl. Acad. Sci. USA* 90:6444-6448 (1993) has provided an alternative mechanism for making bispecific antibody fragments. The fragments comprise a heavy-chain variable domain ( $V_H$ ) connected to a light-chain variable domain ( $V_L$ ) by a linker which is too short to allow pairing between the two domains on the same chain. Accordingly, the  $V_H$  and  $V_L$  domains of one fragment are forced to pair with the complementary  $V_L$  and  $V_H$  domains of another fragment, thereby forming two antigen-binding sites. Another strategy for making bispecific antibody fragments by the use of single-chain Fv (sFv) dimers has also been reported. See, Gruber et al., *J. Immunol.* 152:5368 (1994).

Antibodies with more than two valencies are contemplated. For example, trispecific antibodies can be prepared. Tutt et al., *J. Immunol.* 147:60 (1991). Exemplary bispecific antibodies can bind to two different epitopes, at least one of which originates in the protein antigen of the invention. Alternatively, an anti-antigenic arm of an immunoglobulin molecule can be combined with an arm which binds to a triggering molecule on

a leukocyte such as a T-cell receptor molecule (*e.g.* CD2, CD3, CD28, or B7), or Fc receptors for IgG (Fc $\gamma$ R), such as Fc $\gamma$ RI (CD64), Fc $\gamma$ RII (CD32) and Fc $\gamma$ RIII (CD16) so as to focus cellular defense mechanisms to the cell expressing the particular antigen. Bispecific antibodies can also be used to direct cytotoxic agents to cells which express a particular antigen. These antibodies 5 possess an antigen-binding arm and an arm which binds a cytotoxic agent or a radionuclide chelator, such as EOTUBE, DPTA, DOTA, or TETA. Another bispecific antibody of interest binds the protein antigen described herein and further binds tissue factor (TF).

#### **5.13.6 Heteroconjugate Antibodies**

10 Heteroconjugate antibodies are also within the scope of the present invention. Heteroconjugate antibodies are composed of two covalently joined antibodies. Such antibodies have, for example, been proposed to target immune system cells to unwanted cells (U.S. Patent No. 4,676,980), and for treatment of HIV infection (WO 91/00360; WO 92/200373; EP 03089). It is contemplated that the antibodies can be prepared *in vitro* using known methods in synthetic 15 protein chemistry, including those involving crosslinking agents. For example, immunotoxins can be constructed using a disulfide exchange reaction or by forming a thioether bond. Examples of suitable reagents for this purpose include iminothiolate and methyl-4-mercaptobutyrimidate and those disclosed, for example, in U.S. Patent No. 4,676,980.

20 **5.13.7 Effector Function Engineering**

It can be desirable to modify the antibody of the invention with respect to effector function, so as to enhance, *e.g.*, the effectiveness of the antibody in treating cancer. For example, cysteine residue(s) can be introduced into the Fc region, thereby allowing interchain disulfide bond formation in this region. The homodimeric antibody thus generated can have improved 25 internalization capability and/or increased complement-mediated cell killing and antibody-dependent cellular cytotoxicity (ADCC). See Caron et al., *J. Exp Med.*, 176: 1191-1195 (1992) and Shopes, *J. Immunol.*, 148: 2918-2922 (1992). Homodimeric antibodies with enhanced anti-tumor activity can also be prepared using heterobifunctional cross-linkers as described in Wolff et al. *Cancer Research*, 53: 2560-2565 (1993). Alternatively, an antibody can be engineered that 30 has dual Fc regions and can thereby have enhanced complement lysis and ADCC capabilities. See Stevenson et al., *Anti-Cancer Drug Design*, 3: 219-230 (1989).

#### **5.13.8 Immunoconjugates**

The invention also pertains to immunoconjugates comprising an antibody conjugated to a 35 cytotoxic agent such as a chemotherapeutic agent, toxin (*e.g.*, an enzymatically active toxin of

bacterial, fungal, plant, or animal origin, or fragments thereof), or a radioactive isotope (*i.e.*, a radioconjugate).

Chemotherapeutic agents useful in the generation of such immunoconjugates have been described above. Enzymatically active toxins and fragments thereof that can be used include 5 diphtheria A chain, nonbinding active fragments of diphtheria toxin, exotoxin A chain (from *Pseudomonas aeruginosa*), ricin A chain, abrin A chain, modeccin A chain, alpha-sarcin, Aleurites fordii proteins, dianthin proteins, Phytolaca americana proteins (PAPI, PAPII, and PAP-S), momordica charantia inhibitor, curcin, crotin, sapaonaria officinalis inhibitor, gelonin, mitogellin, restrictocin, phenomycin, enomycin, and the trichothecenes. A variety of 10 radionuclides are available for the production of radioconjugated antibodies. Examples include  $^{212}\text{Bi}$ ,  $^{131}\text{I}$ ,  $^{131}\text{In}$ ,  $^{90}\text{Y}$ , and  $^{186}\text{Re}$ .

Conjugates of the antibody and cytotoxic agent are made using a variety of bifunctional protein-coupling agents such as N-succinimidyl-3-(2-pyridyldithiol) propionate (SPDP), iminothiolane (IT), bifunctional derivatives of imidoesters (such as dimethyl adipimidate HCL), 15 active esters (such as disuccinimidyl suberate), aldehydes (such as glutaraldehyde), bis-azido compounds (such as bis (p-azidobenzoyl) hexanediamine), bis-diazonium derivatives (such as bis-(p-diazoniumbenzoyl)-ethylenediamine), diisocyanates (such as tolyene 2,6-diisocyanate), and bis-active fluorine compounds (such as 1,5-difluoro-2,4-dinitrobenzene). For example, a ricin immunotoxin can be prepared as described in Vitetta et al., *Science*, 238: 1098 (1987). 20 Carbon-14-labeled 1-isothiocyanatobenzyl-3-methyldiethylene triaminepentaacetic acid (MX-DTPA) is an exemplary chelating agent for conjugation of radionucleotide to the antibody. See WO94/11026.

In another embodiment, the antibody can be conjugated to a "receptor" (such as streptavidin) for utilization in tumor pretargeting wherein the antibody-receptor conjugate is 25 administered to the patient, followed by removal of unbound conjugate from the circulation using a clearing agent and then administration of a "ligand" (*e.g.*, avidin) that is in turn conjugated to a cytotoxic agent.

#### 4.14 COMPUTER READABLE SEQUENCES

30 In one application of this embodiment, a nucleotide sequence of the present invention can be recorded on computer readable media. As used herein, "computer readable media" refers to any medium which can be read and accessed directly by a computer. Such media include, but are not limited to: magnetic storage media, such as floppy discs, hard disc storage medium, and magnetic tape; optical storage media such as CD-ROM; electrical storage media such as RAM 35 and ROM; and hybrids of these categories such as magnetic/optical storage media. A skilled

artisan can readily appreciate how any of the presently known computer readable mediums can be used to create a manufacture comprising computer readable medium having recorded thereon a nucleotide sequence of the present invention. As used herein, "recorded" refers to a process for storing information on computer readable medium. A skilled artisan can readily adopt any of the 5 presently known methods for recording information on computer readable medium to generate manufactures comprising the nucleotide sequence information of the present invention.

A variety of data storage structures are available to a skilled artisan for creating a computer readable medium having recorded thereon a nucleotide sequence of the present invention. The choice of the data storage structure will generally be based on the means chosen 10 to access the stored information. In addition, a variety of data processor programs and formats can be used to store the nucleotide sequence information of the present invention on computer readable medium. The sequence information can be represented in a word processing text file, formatted in commercially-available software such as WordPerfect and Microsoft Word, or represented in the form of an ASCII file, stored in a database application, such as DB2, Sybase, 15 Oracle, or the like. A skilled artisan can readily adapt any number of data processor structuring formats (e.g. text file or database) in order to obtain computer readable medium having recorded thereon the nucleotide sequence information of the present invention.

By providing any of the nucleotide sequences SEQ ID NO:1-1009 or a representative fragment thereof; or a nucleotide sequence at least 95% identical to any of the nucleotide 20 sequences of SEQ ID NO:1-1009 in computer readable form, a skilled artisan can routinely access the sequence information for a variety of purposes. Computer software is publicly available which allows a skilled artisan to access sequence information provided in a computer readable medium. The examples which follow demonstrate how software which implements the BLAST (Altschul et al., J. Mol. Biol. 215:403-410 (1990)) and BLAZE (Brutlag et al., Comp. 25 Chem. 17:203-207 (1993)) search algorithms on a Sybase system is used to identify open reading frames (ORFs) within a nucleic acid sequence. Such ORFs may be protein encoding fragments and may be useful in producing commercially important proteins such as enzymes used in fermentation reactions and in the production of commercially useful metabolites.

As used herein, "a computer-based system" refers to the hardware means, software 30 means, and data storage means used to analyze the nucleotide sequence information of the present invention. The minimum hardware means of the computer-based systems of the present invention comprises a central processing unit (CPU), input means, output means, and data storage means. A skilled artisan can readily appreciate that any one of the currently available computer-based systems are suitable for use in the present invention. As stated above, the 35 computer-based systems of the present invention comprise a data storage means having stored

therein a nucleotide sequence of the present invention and the necessary hardware means and software means for supporting and implementing a search means. As used herein, "data storage means" refers to memory which can store nucleotide sequence information of the present invention, or a memory access means which can access manufactures having recorded thereon 5 the nucleotide sequence information of the present invention.

As used herein, "search means" refers to one or more programs which are implemented on the computer-based system to compare a target sequence or target structural motif with the sequence information stored within the data storage means. Search means are used to identify fragments or regions of a known sequence which match a particular target sequence or target 10 motif. A variety of known algorithms are disclosed publicly and a variety of commercially available software for conducting search means are and can be used in the computer-based systems of the present invention. Examples of such software includes, but is not limited to, Smith-Waterman, MacPattern (EMBL), BLASTN and BLASTA (NPOLYPEPTIDEIA). A skilled artisan can readily recognize that any one of the available algorithms or implementing 15 software packages for conducting homology searches can be adapted for use in the present computer-based systems. As used herein, a "target sequence" can be any nucleic acid or amino acid sequence of six or more nucleotides or two or more amino acids. A skilled artisan can readily recognize that the longer a target sequence is, the less likely a target sequence will be present as a random occurrence in the database. The most preferred sequence length of a target 20 sequence is from about 10 to 300 amino acids, more preferably from about 30 to 100 nucleotide residues. However, it is well recognized that searches for commercially important fragments, such as sequence fragments involved in gene expression and protein processing, may be of shorter length.

As used herein, "a target structural motif," or "target motif," refers to any rationally 25 selected sequence or combination of sequences in which the sequence(s) are chosen based on a three-dimensional configuration which is formed upon the folding of the target motif. There are a variety of target motifs known in the art. Protein target motifs include, but are not limited to, enzyme active sites and signal sequences. Nucleic acid target motifs include, but are not limited to, promoter sequences, hairpin structures and inducible expression elements (protein binding 30 sequences).

#### 4.15 TRIPLE HELIX FORMATION

In addition, the fragments of the present invention, as broadly described, can be used to control gene expression through triple helix formation or antisense DNA or RNA, both of which 35 methods are based on the binding of a polynucleotide sequence to DNA or RNA.

Polynucleotides suitable for use in these methods are preferably 20 to 40 bases in length and are designed to be complementary to a region of the gene involved in transcription (triple helix - see Lee et al., Nucl. Acids Res. 6:3073 (1979); Cooney et al., Science 15241:456 (1988); and Dervan et al., Science 251:1360 (1991)) or to the mRNA itself (antisense - Olmno, J. Neurochem. 56:560 (1991); Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988)). Triple helix-formation optimally results in a shut-off of RNA transcription from DNA, while antisense RNA hybridization blocks translation of an mRNA molecule into polypeptide. Both techniques have been demonstrated to be effective in model systems. Information contained in the sequences of the present invention is necessary for the 10 design of an antisense or triple helix oligonucleotide.

#### 4.16 DIAGNOSTIC ASSAYS AND KITS

The present invention further provides methods to identify the presence or expression of one of the ORFs of the present invention, or homolog thereof, in a test sample, using a nucleic acid probe or antibodies of the present invention, optionally conjugated or otherwise associated with a suitable label.

In general, methods for detecting a polynucleotide of the invention can comprise contacting a sample with a compound that binds to and forms a complex with the polynucleotide for a period sufficient to form the complex, and detecting the complex, so that if a complex is 20 detected, a polynucleotide of the invention is detected in the sample. Such methods can also comprise contacting a sample under stringent hybridization conditions with nucleic acid primers that anneal to a polynucleotide of the invention under such conditions, and amplifying annealed polynucleotides, so that if a polynucleotide is amplified, a polynucleotide of the invention is detected in the sample.

In general, methods for detecting a polypeptide of the invention can comprise contacting a sample with a compound that binds to and forms a complex with the polypeptide for a period sufficient to form the complex, and detecting the complex, so that if a complex is detected, a polypeptide of the invention is detected in the sample.

In detail, such methods comprise incubating a test sample with one or more of the 30 antibodies or one or more of the nucleic acid probes of the present invention and assaying for binding of the nucleic acid probes or antibodies to components within the test sample.

Conditions for incubating a nucleic acid probe or antibody with a test sample vary. Incubation conditions depend on the format employed in the assay, the detection methods employed, and the type and nature of the nucleic acid probe or antibody used in the assay. One 35 skilled in the art will recognize that any one of the commonly available hybridization,

amplification or immunological assay formats can readily be adapted to employ the nucleic acid probes or antibodies of the present invention. Examples of such assays can be found in Chard, T., An Introduction to Radioimmunoassay and Related Techniques, Elsevier Science Publishers, Amsterdam, The Netherlands (1986); Bullock, G.R. et al., Techniques in Immunocytochemistry, 5 Academic Press, Orlando, FL Vol. 1 (1982), Vol. 2 (1983), Vol. 3 (1985); Tijssen, P., Practice and Theory of immunoassays: Laboratory Techniques in Biochemistry and Molecular Biology, Elsevier Science Publishers, Amsterdam, The Netherlands (1985). The test samples of the present invention include cells, protein or membrane extracts of cells, or biological fluids such as sputum, blood, serum, plasma, or urine. The test sample used in the above-described method 10 will vary based on the assay format, nature of the detection method and the tissues, cells or extracts used as the sample to be assayed. Methods for preparing protein extracts or membrane extracts of cells are well known in the art and can be readily be adapted in order to obtain a sample which is compatible with the system utilized.

In another embodiment of the present invention, kits are provided which contain the 15 necessary reagents to carry out the assays of the present invention. Specifically, the invention provides a compartment kit to receive, in close confinement, one or more containers which comprises: (a) a first container comprising one of the probes or antibodies of the present invention; and (b) one or more other containers comprising one or more of the following: wash reagents, reagents capable of detecting presence of a bound probe or antibody.

20 In detail, a compartment kit includes any kit in which reagents are contained in separate containers. Such containers include small glass containers, plastic containers or strips of plastic or paper. Such containers allows one to efficiently transfer reagents from one compartment to another compartment such that the samples and reagents are not cross-contaminated, and the agents or solutions of each container can be added in a quantitative fashion from one 25 compartment to another. Such containers will include a container which will accept the test sample, a container which contains the antibodies used in the assay, containers which contain wash reagents (such as phosphate buffered saline, Tris-buffers, etc.), and containers which contain the reagents used to detect the bound antibody or probe. Types of detection reagents include labeled nucleic acid probes, labeled secondary antibodies, or in the alternative, if the 30 primary antibody is labeled, the enzymatic, or antibody binding reagents which are capable of reacting with the labeled antibody. One skilled in the art will readily recognize that the disclosed probes and antibodies of the present invention can be readily incorporated into one of the established kit formats which are well known in the art.

35 **4.17 MEDICAL IMAGING**

The novel polypeptides and binding partners of the invention are useful in medical imaging of sites expressing the molecules of the invention (e.g., where the polypeptide of the invention is involved in the immune response, for imaging sites of inflammation or infection). See, e.g., Kunkel et al., U.S. Pat. NO. 5,413,778. Such methods involve chemical attachment of 5 a labeling or imaging agent, administration of the labeled polypeptide to a subject in a pharmaceutically acceptable carrier, and imaging the labeled polypeptide *in vivo* at the target site.

#### 4.18 SCREENING ASSAYS

10 Using the isolated proteins and polynucleotides of the invention, the present invention further provides methods of obtaining and identifying agents which bind to a polypeptide encoded by an ORF corresponding to any of the nucleotide sequences set forth in SEQ ID NO:1-1009, or bind to a specific domain of the polypeptide encoded by the nucleic acid. In detail, said method comprises the steps of:

15 (a) contacting an agent with an isolated protein encoded by an ORF of the present invention, or nucleic acid of the invention; and  
(b) determining whether the agent binds to said protein or said nucleic acid.

20 In general, therefore, such methods for identifying compounds that bind to a polynucleotide of the invention can comprise contacting a compound with a polynucleotide of the invention for a time sufficient to form a polynucleotide/compound complex, and detecting the complex, so that if a polynucleotide/compound complex is detected, a compound that binds to a polynucleotide of the invention is identified.

25 Likewise, in general, therefore, such methods for identifying compounds that bind to a polypeptide of the invention can comprise contacting a compound with a polypeptide of the invention for a time sufficient to form a polypeptide/compound complex, and detecting the complex, so that if a polypeptide/compound complex is detected, a compound that binds to a polynucleotide of the invention is identified.

30 Methods for identifying compounds that bind to a polypeptide of the invention can also comprise contacting a compound with a polypeptide of the invention in a cell for a time sufficient to form a polypeptide/compound complex, wherein the complex drives expression of a receptor gene sequence in the cell, and detecting the complex by detecting reporter gene sequence expression, so that if a polypeptide/compound complex is detected, a compound that binds a polypeptide of the invention is identified.

35 Compounds identified via such methods can include compounds which modulate the activity of a polypeptide of the invention (that is, increase or decrease its activity, relative to

activity observed in the absence of the compound). Alternatively, compounds identified via such methods can include compounds which modulate the expression of a polynucleotide of the invention (that is, increase or decrease expression relative to expression levels observed in the absence of the compound). Compounds, such as compounds identified via the methods of the 5 invention, can be tested using standard assays well known to those of skill in the art for their ability to modulate activity/expression.

The agents screened in the above assay can be, but are not limited to, peptides, carbohydrates, vitamin derivatives, or other pharmaceutical agents. The agents can be selected and screened at random or rationally selected or designed using protein modeling techniques.

10 For random screening, agents such as peptides, carbohydrates, pharmaceutical agents and the like are selected at random and are assayed for their ability to bind to the protein encoded by the ORF of the present invention. Alternatively, agents may be rationally selected or designed. As used herein, an agent is said to be "rationally selected or designed" when the agent is chosen based on the configuration of the particular protein. For example, one skilled in the art can 15 readily adapt currently available procedures to generate peptides, pharmaceutical agents and the like, capable of binding to a specific peptide sequence, in order to generate rationally designed antipeptide peptides, for example see Hurby et al., *Application of Synthetic Peptides: Antisense Peptides*, In *Synthetic Peptides, A User's Guide*, W.H. Freeman, NY (1992), pp. 289-307, and Kaspczak et al., *Biochemistry* 28:9230-8 (1989), or pharmaceutical agents, or the like.

20 In addition to the foregoing, one class of agents of the present invention, as broadly described, can be used to control gene expression through binding to one of the ORFs or EMFs of the present invention. As described above, such agents can be randomly screened or rationally designed/selected. Targeting the ORF or EMF allows a skilled artisan to design sequence specific or element specific agents, modulating the expression of either a single ORF or 25 multiple ORFs which rely on the same EMF for expression control. One class of DNA binding agents are agents which contain base residues which hybridize or form a triple helix formation by binding to DNA or RNA. Such agents can be based on the classic phosphodiester, ribonucleic acid backbone, or can be a variety of sulfhydryl or polymeric derivatives which have base attachment capacity.

30 Agents suitable for use in these methods preferably contain 20 to 40 bases and are designed to be complementary to a region of the gene involved in transcription (triple helix - see Lee et al., *Nucl. Acids Res.* 6:3073 (1979); Cooney et al., *Science* 241:456 (1988); and Dervan et al., *Science* 251:1360 (1991)) or to the mRNA itself (antisense - Okano, *J. Neurochem.* 56:560 (1991); *Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression*, CRC Press, Boca 35 Raton, FL (1988)). Triple helix-formation optimally results in a shut-off of RNA transcription

from DNA, while antisense RNA hybridization blocks translation of an mRNA molecule into polypeptide. Both techniques have been demonstrated to be effective in model systems.

Information contained in the sequences of the present invention is necessary for the design of an antisense or triple helix oligonucleotide and other DNA binding agents.

5 Agents which bind to a protein encoded by one of the ORFs of the present invention can be used as a diagnostic agent. Agents which bind to a protein encoded by one of the ORFs of the present invention can be formulated using known techniques to generate a pharmaceutical composition.

#### 10 4.19 USE OF NUCLEIC ACIDS AS PROBES

Another aspect of the subject invention is to provide for polypeptide-specific nucleic acid hybridization probes capable of hybridizing with naturally occurring nucleotide sequences. The hybridization probes of the subject invention may be derived from any of the nucleotide sequences SEQ ID NO:1-1009. Because the corresponding gene is only expressed in a limited 15 number of tissues, a hybridization probe derived from of any of the nucleotide sequences SEQ ID NO:1-1009 can be used as an indicator of the presence of RNA of cell type of such a tissue in a sample.

Any suitable hybridization technique can be employed, such as, for example, *in situ* hybridization. PCR as described in US Patents Nos. 4,683,195 and 4,965,188 provides 20 additional uses for oligonucleotides based upon the nucleotide sequences. Such probes used in PCR may be of recombinant origin, may be chemically synthesized, or a mixture of both. The probe will comprise a discrete nucleotide sequence for the detection of identical sequences or a degenerate pool of possible sequences for identification of closely related genomic sequences.

Other means for producing specific hybridization probes for nucleic acids include the 25 cloning of nucleic acid sequences into vectors for the production of mRNA probes. Such vectors are known in the art and are commercially available and may be used to synthesize RNA probes *in vitro* by means of the addition of the appropriate RNA polymerase as T7 or SP6 RNA polymerase and the appropriate radioactively labeled nucleotides. The nucleotide sequences may be used to construct hybridization probes for mapping their respective genomic sequences. The 30 nucleotide sequence provided herein may be mapped to a chromosome or specific regions of a chromosome using well known genetic and/or chromosomal mapping techniques. These techniques include *in situ* hybridization, linkage analysis against known chromosomal markers, hybridization screening with libraries or flow-sorted chromosomal preparations specific to known chromosomes, and the like. The technique of fluorescent *in situ* hybridization of

chromosome spreads has been described, among other places, in Verma et al (1988) Human Chromosomes: A Manual of Basic Techniques, Pergamon Press, New York NY.

Fluorescent *in situ* hybridization of chromosomal preparations and other physical chromosome mapping techniques may be correlated with additional genetic map data. Examples 5 of genetic map data can be found in the 1994 Genome Issue of Science (265:1981f). Correlation between the location of a nucleic acid on a physical chromosomal map and a specific disease (or predisposition to a specific disease) may help delimit the region of DNA associated with that genetic disease. The nucleotide sequences of the subject invention may be used to detect differences in gene sequences between normal, carrier or affected individuals.

#### 10 4.20 PREPARATION OF SUPPORT BOUND OLIGONUCLEOTIDES

Oligonucleotides, *i.e.*, small nucleic acid segments, may be readily prepared by, for example, directly synthesizing the oligonucleotide by chemical means, as is commonly practiced using an automated oligonucleotide synthesizer.

Support bound oligonucleotides may be prepared by any of the methods known to those of 15 skill in the art using any suitable support such as glass, polystyrene or Teflon. One strategy is to precisely spot oligonucleotides synthesized by standard synthesizers. Immobilization can be achieved using passive adsorption (Inouye & Hondo, (1990) *J. Clin. Microbiol.* 28(6) 1469-72); using UV light (Nagata *et al.*, 1985; Dahlen *et al.*, 1987; Morrissey & Collins, (1989) *Mol. Cell Probes* 3(2) 189-207) or by covalent binding of base modified DNA (Keller *et al.*, 1988; 1989); all 20 references being specifically incorporated herein.

Another strategy that may be employed is the use of the strong biotin-streptavidin 25 interaction as a linker. For example, Broude *et al.* (1994) *Proc. Natl. Acad. Sci. USA* 91(8) 3072-6, describe the use of biotinylated probes, although these are duplex probes, that are immobilized on streptavidin-coated magnetic beads. Streptavidin-coated beads may be purchased from Dynal, Oslo. Of course, this same linking chemistry is applicable to coating any surface with streptavidin. Biotinylated probes may be purchased from various sources, such as, *e.g.*, Operon Technologies 30 (Alameda, CA).

Nunc Laboratories (Naperville, IL) is also selling suitable material that could be used. Nunc Laboratories have developed a method by which DNA can be covalently bound to the microwell 30 surface termed Covalink NH. CovaLink NH is a polystyrene surface grafted with secondary amino groups (>NH) that serve as bridge-heads for further covalent coupling. CovaLink Modules may be purchased from Nunc Laboratories. DNA molecules may be bound to CovaLink exclusively at the 5'-end by a phosphoramidate bond, allowing immobilization of more than 1 pmol of DNA (Rasmussen *et al.*, (1991) *Anal. Biochem.* 198(1) 138-42).

The use of CovaLink NH strips for covalent binding of DNA molecules at the 5'-end has been described (Rasmussen et al., (1991). In this technology, a phosphoramidate bond is employed (Chu et al., (1983) Nucleic Acids Res. 11(8) 6513-29). This is beneficial as immobilization using only a single covalent bond is preferred. The phosphoramidate bond joins the DNA to the

5 CovaLink NH secondary amino groups that are positioned at the end of spacer arms covalently grafted onto the polystyrene surface through a 2 nm long spacer arm. To link an oligonucleotide to CovaLink NH via an phosphoramidate bond, the oligonucleotide terminus must have a 5'-end phosphate group. It is, perhaps, even possible for biotin to be covalently bound to CovaLink and then streptavidin used to bind the probes.

10 More specifically, the linkage method includes dissolving DNA in water (7.5 ng/ul) and denaturing for 10 min. at 95°C and cooling on ice for 10 min. Ice-cold 0.1 M 1-methylimidazole, pH 7.0 (1-MeIm<sub>7</sub>), is then added to a final concentration of 10 mM 1-MeIm<sub>7</sub>. A ss DNA solution is then dispensed into CovaLink NH strips (75 ul/well) standing on ice.

15 Carbodiimide 0.2 M 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide (EDC), dissolved in 10 mM 1-MeIm<sub>7</sub>, is made fresh and 25 ul added per well. The strips are incubated for 5 hours at 50°C. After incubation the strips are washed using, e.g., Nunc-Immuno Wash; first the wells are washed 3 times, then they are soaked with washing solution for 5 min., and finally they are washed 3 times (where in the washing solution is 0.4 N NaOH, 0.25% SDS heated to 50°C).

20 It is contemplated that a further suitable method for use with the present invention is that described in PCT Patent Application WO 90/03382 (Southern & Maskos), incorporated herein by reference. This method of preparing an oligonucleotide bound to a support involves attaching a nucleoside 3'-reagent through the phosphate group by a covalent phosphodiester link to aliphatic hydroxyl groups carried by the support. The oligonucleotide is then synthesized on the supported nucleoside and protecting groups removed from the synthetic oligonucleotide chain under standard 25 conditions that do not cleave the oligonucleotide from the support. Suitable reagents include nucleoside phosphoramidite and nucleoside hydrogen phosphorate.

30 An on-chip strategy for the preparation of DNA probe for the preparation of DNA probe arrays may be employed. For example, addressable laser-activated photodeprotection may be employed in the chemical synthesis of oligonucleotides directly on a glass surface, as described by Fodor et al. (1991) Science 251(4995) 767-73, incorporated herein by reference. Probes may also be immobilized on nylon supports as described by Van Ness et al. (1991) Nucleic Acids Res. 19(12) 3345-50; or linked to Teflon using the method of Duncan & Cavalier (1988) Anal. Biochem. 169(1) 104-8; all references being specifically incorporated herein.

To link an oligonucleotide to a nylon support, as described by Van Ness *et al.* (1991), requires activation of the nylon surface via alkylation and selective activation of the 5'-amine of oligonucleotides with cyanuric chloride.

One particular way to prepare support bound oligonucleotides is to utilize the 5 light-generated synthesis described by Pease *et al.*, (1994) PNAS USA 91(11) 5022-6, incorporated herein by reference). These authors used current photolithographic techniques to generate arrays of immobilized oligonucleotide probes (DNA chips). These methods, in which light is used to direct the synthesis of oligonucleotide probes in high-density, miniaturized arrays, utilize photolabile 5'-protected *N*-acyl-deoxynucleoside phosphoramidites, surface linker chemistry and versatile 10 combinatorial synthesis strategies. A matrix of 256 spatially defined oligonucleotide probes may be generated in this manner.

#### 4.21 PREPARATION OF NUCLEIC ACID FRAGMENTS

The nucleic acids may be obtained from any appropriate source, such as cDNAs, genomic DNA, chromosomal DNA, microdissected chromosome bands, cosmid or YAC inserts, and RNA, 15 including mRNA without any amplification steps. For example, Sambrook *et al.* (1989) describes three protocols for the isolation of high molecular weight DNA from mammalian cells (p. 9.14-9.23).

DNA fragments may be prepared as clones in M13, plasmid or lambda vectors and/or 20 prepared directly from genomic DNA or cDNA by PCR or other amplification methods. Samples may be prepared or dispensed in multiwell plates. About 100-1000 ng of DNA samples may be prepared in 2-500 ml of final volume.

The nucleic acids would then be fragmented by any of the methods known to those of skill in the art including, for example, using restriction enzymes as described at 9.24-9.28 of Sambrook *et al.* (1989), shearing by ultrasound and NaOH treatment.

25 Low pressure shearing is also appropriate, as described by Schriefer *et al.* (1990) Nucleic Acids Res. 18(24) 7455-6, incorporated herein by reference). In this method, DNA samples are passed through a small French pressure cell at a variety of low to intermediate pressures. A lever device allows controlled application of low to intermediate pressures to the cell. The results of these studies indicate that low-pressure shearing is a useful alternative to sonic and enzymatic DNA 30 fragmentation methods.

One particularly suitable way for fragmenting DNA is contemplated to be that using the two base recognition endonuclease, *Cvi*JI, described by Fitzgerald *et al.* (1992) Nucleic Acids Res. 20(14) 3753-62. These authors described an approach for the rapid fragmentation and fractionation

of DNA into particular sizes that they contemplated to be suitable for shotgun cloning and sequencing.

The restriction endonuclease *CviJI* normally cleaves the recognition sequence PuGCPy between the G and C to leave blunt ends. Atypical reaction conditions, which alter the specificity of 5 this enzyme (*CviJI*\*\*), yield a quasi-random distribution of DNA fragments from the small molecule pUC19 (2688 base pairs). Fitzgerald *et al.* (1992) quantitatively evaluated the randomness of this fragmentation strategy, using a *CviJI*\*\* digest of pUC19 that was size fractionated by a rapid gel filtration method and directly ligated, without end repair, to a lac Z minus M13 cloning vector. Sequence analysis of 76 clones showed that *CviJI*\*\* restricts pyGCPy and 10 PuGCPu, in addition to PuGCPy sites, and that new sequence data is accumulated at a rate consistent with random fragmentation.

As reported in the literature, advantages of this approach compared to sonication and agarose gel fractionation include: smaller amounts of DNA are required (0.2-0.5 ug instead of 2-5 ug); and fewer steps are involved (no preligation, end repair, chemical extraction, or agarose gel 15 electrophoresis and elution are needed

Irrespective of the manner in which the nucleic acid fragments are obtained or prepared, it is important to denature the DNA to give single stranded pieces available for hybridization. This is achieved by incubating the DNA solution for 2-5 minutes at 80-90°C. The solution is then cooled quickly to 2°C to prevent renaturation of the DNA fragments before they are contacted with the 20 chip. Phosphate groups must also be removed from genomic DNA by methods known in the art.

#### **4.22 PREPARATION OF DNA ARRAYS**

Arrays may be prepared by spotting DNA samples on a support such as a nylon membrane. Spotting may be performed by using arrays of metal pins (the positions of which correspond to an array of wells in a microtiter plate) to repeated by transfer of about 20 nl of a DNA solution to a 25 nylon membrane. By offset printing, a density of dots higher than the density of the wells is achieved. One to 25 dots may be accommodated in 1 mm<sup>2</sup>, depending on the type of label used. By avoiding spotting in some preselected number of rows and columns, separate subsets (subarrays) may be formed. Samples in one subarray may be the same genomic segment of DNA (or the same gene) from different individuals, or may be different, overlapped genomic clones. Each of the 30 subarrays may represent replica spotting of the same samples. In one example, a selected gene segment may be amplified from 64 patients. For each patient, the amplified gene segment may be in one 96-well plate (all 96 wells containing the same sample). A plate for each of the 64 patients is prepared. By using a 96-pin device, all samples may be spotted on one 8 x 12 cm membrane.

Subarrays may contain 64 samples, one from each patient. Where the 96 subarrays are identical, the dot span may be 1 mm<sup>2</sup> and there may be a 1 mm space between subarrays.

Another approach is to use membranes or plates (available from NUNC, Naperville, Illinois) which may be partitioned by physical spacers *e.g.* a plastic grid molded over the membrane, the grid 5 being similar to the sort of membrane applied to the bottom of multiwell plates, or hydrophobic strips. A fixed physical spacer is not preferred for imaging by exposure to flat phosphor-storage screens or x-ray films.

The present invention is illustrated in the following examples. Upon consideration of the present disclosure, one of skill in the art will appreciate that many other embodiments and variations 10 may be made in the scope of the present invention. Accordingly, it is intended that the broader aspects of the present invention not be limited to the disclosure of the following examples. The present invention is not to be limited in scope by the exemplified embodiments which are intended as illustrations of single aspects of the invention, and compositions and methods which are functionally equivalent are within the scope of the invention. Indeed, numerous modifications and 15 variations in the practice of the invention are expected to occur to those skilled in the art upon consideration of the present preferred embodiments. Consequently, the only limitations which should be placed upon the scope of the invention are those which appear in the appended claims.

All references cited within the body of the instant specification are hereby incorporated by reference in their entirety.

20           **5.0       EXAMPLES**

**5.1       EXAMPLE 1**

**Novel Nucleic Acid Sequences Obtained From Various Libraries**

A plurality of novel nucleic acids were obtained from cDNA libraries prepared from various human tissues and in some cases isolated from a genomic library derived from human chromosome 25 using standard PCR, SBH sequence signature analysis and Sanger sequencing techniques. The inserts of the library were amplified with PCR using primers specific for the vector sequences which flank the inserts. Clones from cDNA libraries were spotted on nylon membrane filters and screened with oligonucleotide probes (*e.g.*, 7-mers) to obtain signature sequences. The clones were clustered into groups of similar or identical sequences. Representative clones were selected for 30 sequencing.

In some cases, the 5' sequence of the amplified inserts was then deduced using a typical Sanger sequencing protocol. PCR products were purified and subjected to fluorescent dye terminator cycle sequencing. Single pass gel sequencing was done using a 377 Applied Biosystems

(ABI) sequencer to obtain the novel nucleic acid sequences. In some cases RACE (Random Amplification of cDNA Ends) was performed to further extend the sequence in the 5' direction.

## 5.2 EXAMPLE 2

### 5 Novel Contigs

The novel contigs of the invention were assembled from sequences that were obtained from a cDNA library by methods described in Example 1 above, and in some cases sequences obtained from one or more public databases. Chromatograms were base called and assembled using a software suite from University of Washington, Seattle containing three applications designated 10 PHRED, PHRAP, and CONSED. The sequences for the resulting nucleic acid contigs are designated as SEQ ID NO: 1-1009 and are provided in the attached Sequence Listing. The contigs were assembled using an EST sequence as a seed. Then a recursive algorithm was used to extend the seed EST into an extended assemblage, by pulling additional sequences from different databases (i.e., Hyseq's database containing EST sequences, dbEST version 114, gb pri 114, and UniGene 15 version 101) that belong to this assemblage. The algorithm terminated when there was no additional sequences from the above databases that would extend the assemblage. Inclusion of component sequences into the assemblage was based on a BLASTN hit to the extending assemblage with BLAST score greater than 300 and percent identity greater than 95%.

The nucleotide sequence within the assembled contigs that codes for signal peptide 20 sequences and their cleavage sites was determined from using Neural Network SignalP V1.1 program (from Center for Biological Sequence Analysis, The Technical University of Denmark). The process for identifying prokaryotic and eukaryotic signal peptides and their cleavage sites are also disclosed by Henrik Nielson, Jacob Engelbrecht, Soren Brunak, and Gunnar von Heijne in the publication "Identification of prokaryotic and eukaryotic signal peptides and prediction of their 25 cleavage sites" Protein Engineering, vol. 10, no. 1, pp.1-6 (1997) incorporated herein by reference. A maximum S score and a mean S score, as described in the Nielson et al. reference, are obtained from each assembled contig. Table 3 sets forth the nucleotide range for each sequence of SEQ ID NO: 1-1009 that encodes a corresponding amino acid sequence containing the signal peptide 30 sequence and its cleavage site: the maximum S score and the mean S score obtained for each sequence.

A signal peptide or leader peptide is usually a segment of about 15 to 30 amino acids at the N terminus of protein that enables the protein to be targeted to a cell membrane or secreted from a cell. Generally, the signal peptide acts as an exportable and is removed as the protein is secreted in its final form.

The nearest neighbor result for the assembled contig was obtained by a BLASTX version 2.01al 19 MP-Washington University search against Genpept release 120 and Geneseq database (October 12, 2000, update 21 (Derwent)), using BLAST algorithm. The nearest neighbor result showed the closest homologue for each assemblage from Genpept (and contains the translated 5 amino acid sequences for which the assemblage encodes). The nearest neighbor results for SEQ ID NO: 1-1009 are shown in Table 2.

Tables 1, 2 and 3 follow. Table 1 shows the various tissue sources of SEQ ID NO: 1-1009. Table 2 shows the nearest neighbor result for the assembled contig. The nearest neighbor result shows the closest homolog with an identifiable function for each assemblage. Table 3 contains the 10 start and stop nucleotides for the translated amino acid sequence for which each assemblage encodes. Table 3 also provides a correlation between the amino acid sequences set forth in the Sequence Listing, the nucleotide sequences set forth in the Sequence Listing and the SEQ ID NO. in USSN 09/491,404.

TABLE 1

TISSUE ORIGIN	RNA SOURCE	HYSEQ LIBRARY NAME	SEQ ID NOS: OF NUCLEOTIDE(S)
adult brain	GIBCO	AB3001	31 45 61 78 96 122 126 132 163 169 171-172 175-176 181 203 212 220 222 230 251-252 258 263 267 279 336 343 358 396 400-401 422 428-429 431 437 456 464 487 503 513 524 561 580 583 609 619 682 812 946 958 965 980 983 989 999
adult brain	GIBCO	ABD003	5 23 26 28-29 31 34-36 61 74 78 87 111-113 116 122-123 129 139 143 148 159 163 167 175-176 178 181 183 186 201-204 206 208-209 212 214 220 222 228 230 234-235 237 246 249-250 252 255 259 262- 264 266-267 279-280 286 329 336 351 358 379 396 422 429 431 437 439 444-445 450 452 456 467-468 479 484 503-504 507 513 523-524 526 533 550 553 559 561-562 578 580 583 636 638 640 683 711 759 764 769 772 799 803 824 830 842 865 885 900 902 906 910 922-924 932-933 941 945 951 955 958 965 971 983-984 989 999 1005
adult brain	Clontech	ABR001	81 122 148 181 183 204 207 233 237 250 267 301 346 394 396 437 439 457 505 563 618 653 655 721 764 795 885 942 949
adult brain	Clontech	ABR006	148 152 222 257 269 583 640 677 878
adult brain	Clontech	ABR008	2 10-11 13-14 19-20 23 28-29 34- 35 37 39-40 45 49-50 52 60 73-74 78 83 87-91 94 98 101 109 114-117 122-123 143 145 148-150 152 156 162 168 173-178 181 183 187 189 194 204 206-209 212 214-215 220- 221 228 231 233-238 246-247 249- 253 255-260 262 266 269-270 272 276 278-281 284 294 301 313 316- 320 335 337-338 343 363 372 379 388 390-392 396 400-401 403 405- 407 414 417 422-423 425 427-428 433 437 441 443-446 452-453 456 464 467 469 473-479 482 484 487- 488 491 497-498 500 502 504-505 507 519-520 523-526 533 544-545 553 555-556 563 570-571 574-576 578-580 583 615 618-619 637-638 643-644 653 655-656 661 663 678 680 689-690 695 699 702 705 717- 718 720 722 725-726 742 746 752 754-755 759 761 763-765 767 769 772-774 776 784-789 792 795 799 809-810 812 814-815 817 834 840 842 844-846 852 855-856 858-860 870-873 875 877 885-886 888 890- 897 903-904 910 928 930-932 939- 942 946-947 951-952 955 957 960 964-965 967 971 975-976 978 986- 987 989 992 999 1001
adult brain	Clontech	ABR011	214 965

TABLE 1

TISSUE ORIGIN	RNA SOURCE	HYSEQ LIBRARY NAME	SEQ ID NOS: OF NUCLEOTIDE(S)
adult brain	BioChain	ABR012	152 498
adult brain	Invitrogen	ABR013	142 207 254 396 442 498
adult brain	Invitrogen	ABT004	2 23 31 34 78 96 116 129 141 160 176-177 181 183 202 214 231 233 248 256 258-260 262 278 310 336- 337 379 416 437 439 443-444 450 452 454 464 467 479 484 500 504 519 526 553 570 590 619 638 640 647 653 655 678 711 759 764 789 795 799 885 887 892 902 905 907 910 915 922 941-942 955 960 989 999
cultured preadipocytes	Strategene	ADP001	17 37 39 74 79 111 129 152 160 200 222 248 252 268 274 358 385 450 456 504 526 571 583 619 633 640 740 803 816 829 842 887 939- 940 965 973 977 986
adrenal gland	Clontech	ADR002	4 6 19 36 39 49 51-53 74 76 118 122-123 147-148 152 156 160 167 171-172 181 183 204 206 212 223- 224 228 233-234 246 249-250 254- 255 262 274 278-279 284 287 294 317 336 355 358 366 379 392 401- 402 412 417 420 431-432 439 464 470 479-480 484 503-504 506 509 519 524 526-527 541 553 555 561 583 614 619 631 638 646 682 738- 739 756 760 764 770 800 802-803 816-817 838 847 852 863 881 887 905-906 910 923 926 932 941 950- 951 989 999 1002
adult heart	GIBCO	AHR001	6 20 26 29 31 34 37 39 41 46 61 74 78 101 114 116-118 122-124 128 145 147-148 152 155 163 175-176 178 181 183 200 204 206 210 212 215 228 230 234-235 237 246 248- 252 255-256 262-263 266-268 272 278 280 282-283 286 294 309 313 350-351 358 370 374 379 391-392 394 397 400-401 409 420 423 431- 432 434 436 438 441 443 452 455- 456 461 467-468 479-480 484 487 498 500 503 505 511 519 533 541 550 552-553 558 561-562 568 575 583 590 597-598 603 619 636-638 644-645 667-668 680 684 711-712 714-715 723 732 750 789 803 805 816 822 828 885 889 900 902 905 908 910 916-917 923-924 932 935 937 939 941 950 952 954 960 965 974 982 984 987 993 1005
adult kidney	GIBCO	AKD001	4 13-14 19-20 23 26-31 37 39 47 49 54 61 64 78 81 87 91 98 101 114 118 122-123 127 129-130 141- 143 145 148-149 155-158 160 163 168 171-172 175-176 178-181 183 197-198 200 203-206 208 212 215 221-222 228 230 234 237 241 245- 246 250-252 254-257 262-263 265- 269 278-279 282-284 286 297 301

TABLE 1

TISSUE ORIGIN	RNA SOURCE	HYSEQ LIBRARY NAME	SEQ ID NOS: OF NUCLEOTIDE(S)
			308 333 336 352-353 358 371-372 379 381 386 391 394 396-397 400- 401 405 409 417 420 428-429 431 436-437 443 445 450 456 463-466 468 475 479-480 484 487 495 498- 499 503-505 507 511 513 517 523 526 529 533 539 541-542 550 552- 553 555 561 570-572 575 577-578 583 587 597 604 606 609 619 636 638 640-642 648 680 682 701 706 714 721 732 740 747 771 792 803 805 809 811-812 829 838 842 862 865 885 889 900 902 905-906 908 910-911 918-921 924 926 928-930 937 939 941-942 950-951 953 955 958 960 963 965 967 976 978-979 982-984 1005
adult kidney	Invitrogen	AKT002	19 31 78 81 91 98-99 122 142 145 148 152 158 169 176 248 254 256 262 266 279 296-297 301 321 353 372 401 405 416 420 429-430 441 456 464 498 504 507 523 526 533 541 583 592-597 649 701 791 838 862 868 911 926 933 946-947 958 960 971
adult lung	GIBCO	ALG001	19 33 48 61 96 98 101 108 111 114 145 148 179 183 194 198 200 205 212 220 228 234 246 248 250-251 254-255 263 268 277 279 289 298 306 337 343 372 379-380 385 401 405-406 408 410 420 431 440 443 445 449 455 484 499 503 507 513 517 571 590 597 617 636 640 714 732 749-750 805 885 900 905 910 918 941 955 958 960 977 980 1001 1005
lymph node	Clontech	ALN001	43 48 53 108 123 136 142 147 160 178 181 183 200 205 228 244 246 250 254 268 270 291 379 399 419 431 440 442 479-480 484 519 533 539 553 559 565 583 616-617 619 636 662 701 740 805 833 910 913 928 941 977
young liver	GIBCO	ALV001	19 42 45 61 64 84 98 107 109 122- 123 129-130 133 142 148 168-169 178 181 183 200 205 207 227-229 232 238 246-248 250 253-255 262- 263 265 268 279 317 336 371 377 392 400 410 431 436-437 443 445 448-450 484 487 513 533 545 559 561 570 578 617 632 638 640 648 680 771 803 816 836-838 885 906 926 940 986
adult liver	Invitrogen	ALV002	13-14 26 36 54 64 74 76 109 117 122 179 181 183 187 204 215 221 225 229 232 247-248 250 256-257 275 304 307 315 317 321-322 371 377 379 386 416 420 448-449 457 464 475 479 481 483-484 504 507 526 553 557 570 619 627-629 632

TABLE 1

TISSUE ORIGIN	RNA SOURCE	HYSEQ LIBRARY NAME	SEQ ID NOS: OF NUCLEOTIDE(S)
			638 640 653 655 675 680 701 752 768 827 848 865 882 885 889 910 951 955 959 963 967 978 989 999- 1000
adult ovary	Invitrogen	AOV001	4 12 19 23 28-32 34-37 39 45 48 52 54 60-61 64-65 67 76 78 87 96 98-100 108 111-112 114 116-118 122-123 126 129-130 132-134 137 139 142-145 147-149 152 162-163 169-172 176 178 180-183 187 191- 192 197-202 204-206 212 214-217 219-222 228 234-235 237 242 246- 248 250-252 254-256 262 265-269 274 279-280 282-284 294 308-309 313 317 336-337 346 358 361 364 371 374 379 391-392 394 396-397 400 408 414 418 420 423 425 428- 429 431 435-437 440-441 443-447 450 452 455-459 463-464 467-468 479-480 484 487 492 495 499-500 503 505 512-513 517 519 524 533 539 545 553 555 557-559 561 565- 566 568 571 575 577-578 581 583 590 597 605 610 613 616-617 619 636 638 640 645-646 649-650 654 662 671 680 682 694 697 701 711 732 735 739-741 750 753 760 764 771 780 785 789 792 803 806 810 812 821 831-832 838 841-842 879 885 887 900 902 905-906 908-912 917 921-922 924 928 936-939 941- 942 946 950-952 957-958 960 962- 965 979 982 987 989 994 998-999 1005 1008
adult placenta	Clontech	APL001	122 148 168 181 194 200 248 262 268 317 436 541 561 803 838 911 971
placenta	Invitrogen	APL002	38 61 78-79 142 149 176 187 194 206 215 246 252 278 337 346 379 400 456 464 478-479 484 487 504 519 526 553 571 638 640 732 842 910-911 918 941 958
adult spleen	GIBCO	ASP001	23 26 39 43 48 61 63 78 87 98 108 110 123 136 142 157 176 178 181 183 197-198 201-202 205-206 213 220 222 228 234 237 244 250-252 254-255 257 263 294 305 320 336- 337 354 358 371-372 376 379 397 400 405 410 414 431 437 440 455- 456 484 487 498-499 504 506-507 511-512 519 523 526 529 533 539 550 561 565 572 575 583 586 597 616-617 619 621 636 640 687 701 713 732 740 748 803 812 816 835 910 930 939 946 956 958
testis	GIBCO	ATS001	20 23 29 61 64 76 114 123 126 143 145 148-149 175 178 182 200 203 206 209 235 248 252 257 263 268 279-281 283-284 333 358 371 391 396 400 418 423 431 438-439 441

TABLE 1

TISSUE ORIGIN	RNA SOURCE	HYSEQ LIBRARY NAME	SEQ ID NOS: OF NUCLEOTIDE(S)
			445 456 479-480 487 490 505 507-508 516-517 521 524 533 550 559 561-562 582 597 606 638 646 676 680 750 772 803 834 877 908 911 914 937-938 950 989 999
adult bladder	Invitrogen	BLD001	23 37 77-78 84 160 176 178 181 215 218 248 252 262 274 299 334 351 401 464 474 484 517 543 619 663 692 729 908 910 918 937 941 951 960 962
bone marrow	Clontech	BMD001	19 31 39 43 48 52-53 95-96 98 100 108 111-112 114 117 122-123 136 141-142 144-145 147-149 152 161 163 169 181 183 187 194 201 204-205 208 213 222 228 234 241-242 244-246 248-251 254-255 257 267 272 274 282 286 288-289 292 294 313 317 335 337 339 346-347 358 363 365 374 379 391-392 395-398 406 408 414 418 423 428 436 440-442 444-445 456 475 479 484 495 498-500 504 508 511 516 519 526 533 539 541 553 556 559 561 565 571 573 583 597 612 617 619 638 640 646 649 651 677 681 685 707 709-710 721 734 764 771 803 806 811 838 852 858 869 885 908 910 916 922 930 936-937 941 951 965 982 985 989 991 995 999 1005 1008
bone marrow	Clontech	BMD002	31 39 43 48 68 71 91 108 122-123 134 136 142 148-150 152 161 169 178 181 194 196 204-205 208 244 246 254 262-263 265 267 272-273 300 320 343 356 363 372 379 405 408 413-414 430-431 436 440-441 454 479 484 486 512-513 517 519 533 553 559 570 583 590 617-619 634 637 651 674 692 793-794 800 803 818 852 880 904 910 930 936 941 950
bone marrow	Clontech	BMD004	142 152 254 274
adult colon	Invitrogen	CLN001	26 29 48 61 108-109 129-130 144 176 194 215 221 252 401 436 440 450 498 511 533 583 590 616-617 706 764 905 939 955
adult cervix	BioChain	CVX001	6 16 19-20 29 35 37 43 45 64 73 75-76 86 92 96-98 100-101 105 108 111 113 122 143 145 147-149 163-165 167 172 174 178 181-183 187 200-201 206 222 234 237-238 242-243 246 248 250-251 253 261-262 265 268 270 274 279 283-284 294 308 343 345 352 365 379 381 391 400 409 420 423-424 428 436 443-444 463-464 473 479-480 484 487 505 508 510-512 516-517 519 523-524 533 539 553-555 558-559 561-562 575 578 583 591 597 619 643 645-646 650 657 671 680 740 764 771 796 803 811 816 865 889 908

TABLE 1

TISSUE ORIGIN	RNA SOURCE	HYSEQ LIBRARY NAME	SEQ ID NOS: OF NUCLEOTIDE(S)
			910 926-927 933 937 941 960 963 965 967-968 977 982 989 999 1008-1009
diaphragm	BioChain	DIA002	26 152 499 680
endothelial cells	Strategene	EDT001	13-14 19 23 26 30-32 34 39 67 73-74 76 78 91 101 109 114 116 118 129 145 149 152 156 160-161 167 176 180 183 187 197 201 203-204 206 209 215 222 226 228 230 237 246 248 250-252 256-257 262 266 276 279 282-283 286 309 312-313 343 358 372 391-392 394 396 400-401 405 409 413 420 423 429-431 436 438 443-445 450 455-456 479 484 487 498-499 503 507 509 511 513 523 561-562 571 575 583 619 639 646 653 655 680 711 721 729 739 771-772 775 779 795 803 805 834 838-840 885 889 900 905-906 911 917-918 922 924 930 942 946 955 958 960 977-979 982-984
Genomic clones from the short arm of chromosome 8	Genomic DNA from Genetic Research	EPM001	122 148 436
Genomic clones from the short arm of chromosome 8	Genomic DNA from Genetic Research	EPM003	122 148 379 436
Genomic clones from the short arm of chromosome 8	Genomic DNA from Genetic Research	EPM004	122 148 436
Genomic clones from the short arm of chromosome 8	Genomic DNA from Genetic Research	EPM005	148
esophagus	BioChain	ESO002	152 178 583
fetal brain	Clontech	FBR001	122 148 181 279 284 484 553 575 619 668 911
fetal brain	Clontech	FBR004	122 190 212 379 479 484 541 905 922 924 941 950
fetal brain	Clontech	FBR006	2 23 31 36 39 42 44 49 52 78 87 114 117 122-123 145 148 176-177 180-181 187 204 208 210 215 220 235 238-239 241 245-246 251 253 256 259 266 270 278 280 286 314 317 337 372 379 392 396 400-401 405-406 410 414 423 428 439-440 443 445 452 467 473 479 484 487 491 497 500 504 517 519 524 526 544 553 556 561 563 568 570-571 573 577 586 619 647 653 655 664-665 680 739 742 746 754 766 772-776 784 795 798 834 840 842 863 878 885 892-893 898-899 910 930 941-942 946 952 965 971 976 987 993
fetal brain	Invitrogen	FBT002	19 31 34-35 44-45 78-79 87 96 101 116 129 176 181 204 206 233 235

TABLE 1

TISSUE ORIGIN	RNA SOURCE	HYSEQ LIBRARY NAME	SEQ ID NOS: OF NUCLEOTIDE(S)
			256-257 259 262 278 280 317 320 337 380 396-397 401 437 443 446 450 453 464 480 484 498-499 504 526 577 591 619 640 664 680 697 710 764 900 902 905 910 958
fetal heart	Invitrogen	FHR001	500 910
fetal kidney	Clontech	FKD001	39 47 96 98 122-123 148 156 181 200 207 246 268 274 279 283 300 379 411 445 464 468 479 484 506 542 553 561 583 619 680 686 712 747 910 941
fetal kidney	Clontech	FKD002	479 484 583 803 910 941
fetal kidney	Invitrogen	FKD007	864
fetal lung	Clontech	FLG001	64 96 143-144 168 194 206 234 266 335 337 363 500 507 561 619 968
fetal lung	Invitrogen	FLG003	3 13-14 55 61 79 122-123 148 160 181 183 194 200 234 248 250 252 266 268 273 289 294 336 358 428 432 436 484 507 510 513-514 533 541 557-558 582-583 597 671 711 764 777 806 811 817 905 933 978
fetal lung	Clontech	FLG004	951
fetal liver-spleen	Columbia University	FLS001	13-15 19-21 23-26 28-30 32 34 37 39 45 47-49 56 67 72-74 78 84 87 91 96-98 101 103-104 108 111 114 116 122-123 126 129 131 133 142- 145 147-149 151-152 156 160-161 166 168-169 172 176 178-179 181 183-185 192-194 197-202 204-206 208 215 221-222 224 228-229 232 234-235 237 246 248-252 254-257 262 266-268 272 274 278-280 282- 287 294 313 315 321 333 336-337 343-344 358 372 377-379 386 391- 393 397 400-402 404-405 409-410 418 420-421 429 431 436-437 440- 441 443 445 448-450 456-457 464 473 475 478-481 483-484 487-488 498 500 503 505 507 509 513 522- 523 528 533-534 541 551 553 558 560-562 564-565 570 575 577-578 583 586 590 597 600 605-607 617 619 632 636 638 640 644 646 672 677-680 705 711 729 732 735-738 740 742 748 760 763-764 771-772 792 802-803 805-806 812 816-817 820-821 824-827 834 838 842-843 848 853 861 865 878 885 887 889 900 902 904-906 908 910-911 917 924 926 928 930 934 936-937 941 944 946 950-951 955 958 960 963 965 974-980 982-983 988-990 999
fetal liver-spleen	Columbia University	FLS002	4 8 12 15-16 18-21 23-24 26 32 37 39 47 54 61 64 67 71-72 74 76 79 83-84 87 91 96-98 100-104 109 111-113 122-123 129 133 141 145 147-149 152 161 163 169 171-172 174 178-181 183 185 187-188 192- 195 198-202 205 207-209 213 215 221-222 229 232 234-235 237 241

TABLE 1

TISSUE ORIGIN	RNA SOURCE	HYSEQ LIBRARY NAME	SEQ ID NOS: OF NUCLEOTIDE(S)
			244-246 248 250 262 265 267-268 270 274 278-280 283-284 290 294 300 311 313-315 317 331 337 341 346 351-352 358 360-361 371-372 377 382 391-393 397 399-401 404- 405 410 414 425 429 431 436 440- 441 445-446 448-450 453 456 464 473 475 479-480 487 492 498 500 503-504 507 512 517 519 523 526 540 557 561-563 565 574-575 577- 578 583 590 597 605-606 608 611 614 616 619 631-634 636-638 640 646 649-650 662 671-673 676-678 682 684 701-702 704-705 711 716 732 735 748 760 762-764 768 771- 772 779 790 802 805 815-816 834 838 842 848 865 878-879 883 887- 889 903 905-906 910 916-917 922 924 928 930 939 944 946 950 955- 956 958 960 965 975 977 982-983 987-988 993-994 998 1004
fetal liver-spleen	Columbia University	FLS003	377 732 889 938
fetal liver	Invitrogen	FLV001	23 29 39 84 109 194 208 221 232 247-248 278 301 321 336-337 370- 371 379 443 448-449 464 475 479- 480 498 500 533 550 578 590 632 636 640 678 680 683 751 763 803 882-883 885 887-889 910 921 942 946 951 963 988
fetal liver	Clontech	FLV004	37 122 200 232 268 274 377 583 946
fetal muscle	Invitrogen	FMS001	29 37 41 64 66 74 148 164 200 202 208-209 252 257 259 262 265 268 274 279 337 346 379 445 480-481 505 507 553 555 561 571 606 640 676 781 801 838 910 926 928 951 957 960 963 965
fetal muscle	Invitrogen	FMS002	200 268 274
fetal skin	Invitrogen	FSK001	23 29 31 34 49 78 84 87 96 100 112 116 133 143 148 163 168 172 176-177 181 193 199-202 208 215 222 235 240 246 248 252 256-257 262-268 274 280 282 294 309 314 317 322 346 358 371 373-375 379 414 417 419-420 436-437 441 445 454 456 458 479-480 484 499-500 504 507 513 519-520 526 533 539 541 545-547 550 561 565 570-571 575 577 583 590 598-599 619 644 650 665 697 702 706 739 742 744 784 790 792-793 812 816 861 877 889 906 910 918 922 941 949 951- 952 955 962 964-965 968 979 983 987 989 999
fetal skin	Invitrogen	FSK002	200 257 265 268 274 513 688
fetal spleen	BioChain	FSP001	39 431 523 533 617
umbilical cord	BioChain	FUC001	19 28-29 34 39 74 96 99 101 111 114 116 122 143 145 148 163 168 175 178 181 183 197 200 205 212

TABLE 1

TISSUE ORIGIN	RNA SOURCE	HYSEQ LIBRARY NAME	SEQ ID NOS: OF NUCLEOTIDE(S)
			222 228 230 237-238 246 248 252-253 255 257 259 262 265 268-269 272 274 282 325 351 379 396 400-401 413 429 441 443 445 452 456-457 467-468 479 484 487 505 513 517 519 523 533 541 553 555 561 571 575 577 583 590 601-602 605-606 619 636 645 680 693 698 711 757 759 764 803 814 816 821 853 885 889 900 906 908 910 924 926 932 937 941 943 946 951-952 955 958 976 987 989 993-994 999
fetal brain	GIBCO	HFB001	13-14 19 26 29 31-32 39 44-45 61 67 74 78 88 100 114 122-123 126 129 148 152 163 167 169 171-172 175-176 180-181 187 201-204 206 209 212 215 220 222 227-228 230 233-235 237 246 249 251 258-259 262-263 266 269 279-280 282 284 286 333 337 340 342 355 358 362 366 379 391 394-397 406 422-423 428-429 431 436-437 443-446 450 452 456 467-468 479-480 484 498 504-505 513 517 523 526-527 533 539 541 558-559 561-562 574 580 583 605 619 635 638 643 680 682 708 711 739-740 742 764 776 803 812 823 865 885 900 902 905 910 917 924 928 932 939 941 945 958 960 964-965 974 978-979 984
macrophage	Invitrogen	HMP001	152 201 498 983
infant brain	Columbia University	IB2002	2 20 23 26 28-29 31 37 39 44 57 74 78-79 111 118 122-123 126 129 143 145 148 155 168-169 175-176 178 181 185-186 191 200-202 208 212 214-215 220 222 224 228 230-231 235 237 239 248-249 252 255-260 262 266-269 272 280 284 286 289 313 323 326 329 346 358 361 379 396 400 412 422-423 428 437 439 443 445 450 452 457 461 467-468 479-480 484 487 490 498 500 504-505 523 526 533 541-542 547 561-562 571 574-575 580 605 635 637 640 647 653 655 678 680 711 733 746 761 764 766 771 776 795 865 885 887 900-901 905 907 910 917 924 930 932 941-942 951 958 960 962 967 974-975 979 982-983 989 993 999 1003-1004
infant brain	Columbia University	IB2003	23 31 53 87 107 123 160 175 185 197 202 207 215 222 237 252 256-258 274 284 289 326 358 396 400 437 445 452 462 464 467 487 500 504 526 575 583 590 605 630 653 655 703 733 757 764 795 865 884-885 900 905 919 924 974-975 982
infant brain	Columbia University	IBM002	44 169 248 746 764 958
infant brain	Columbia	IBS001	76 119 147 149 181 248 329 361

TABLE 1

TISSUE ORIGIN	RNA SOURCE	HYSEQ LIBRARY NAME	SEQ ID NOS: OF NUCLEOTIDE(S)
	University		379 764 910 942 951
lung, fibroblast	Strategene	LFB001	13-14 26 78 84 91 98 114 122 148 176 197 204 222 246 251 266 379 387 431 437 441 464 479 484 533 553 571 583 619 645-646 711 739 752 910 926 950 965 978 984
lung tumor	Invitrogen	LGT002	13-14 19 31-32 34-39 43 48 64 67 74 76 87 93 95-96 101 111-112 116 122-123 134 138 142 144-145 147- 148 151-152 160 172 178-179 181- 183 187 191-194 197-198 200-202 205 208 210 218 226 228 234 237 246 248 250-252 254-255 257 260- 262 265 268 274 277-279 289 301 320-321 333 336 343 352 355 358 366-368 371 374 379 391-392 397 400-401 406 410 414 423 431 436 440-441 455-456 458 463-464 468 478-480 484 487 498 503-504 511 519 526-527 529 533 541 553 557 561 570-571 575 578 581 583-586 588-589 597 606 616 619 636 638 640 648 650 652 657 680 700 705- 706 708 716 721-722 729 732 739 744-745 752 762 764 782 795 803 812 816-817 838 863 874 877 906 910-911 922 926 941 951 955 957- 958 962-963 968-969 977-978 982- 983 996-997 1007
lymphocytes	ATCC	LPC001	13-14 35 66 79 95 106-107 112 122-123 149 152 178 181 201 205 246 251-252 267 293 299 358 379 384 400-401 409 415 418 439 443- 444 451 456 458 479 484 487 513 533 568 572 575 583 614 619 686 706 721 730-731 739 747 764 789 905 910 941-942 950 965 978-979 1007
leukocyte	GIBCO	LUC001	13-14 19 23 30-32 36 39 45 48-49 60-61 63 67 73-74 78-79 81-82 84 87 91 98-99 107-109 111-112 114 122-123 129 142 144-145 148-150 152 170 176 179 181 183 187-188 194 198 201-208 212-213 215 222 228 235 237 241-242 244-246 249- 251 254-257 263 267 278-280 282- 284 286 289-290 295 302 308-309 313 317 333 337 343 346 356-358 371 379 391-392 394 397 400-401 404 406-410 412-415 423-424 429 431 436 439-441 443-445 450 456 458 479-480 484 487-488 495 498- 500 503 505 511-514 519 523 530- 533 539 541 555 559 561 565-566 570 572 577-578 583 590 595 597 617 619 633 635-636 639-640 646 660 670 672 677 680-681 698 703 705 729 732 739-740 743 747 750 763-764 771 782 792-793 803-805 809 819 838 857 866-867 885 888

TABLE 1

TISSUE ORIGIN	RNA SOURCE	HYSEQ LIBRARY NAME	SEQ ID NOS: OF NUCLEOTIDE(S)
			900 905 910-911 924 926 928 930 941 948 950-953 955 962-963 965 977-979 984 987 989 999 1008
leukocyte	Clontech	LUC003	19 26 68 76 96 122 147 152 198 201 205 208 284 317 354 358 430 436 440 479 511 533 541 553 561 583 589 646 698 732 764 766 838 984
melanoma from cell line ATCC #CRL 1424	Clontech	MEL004	8 23 36 69 91 114 122-123 126 148 151 181 202 204 227 246 256-257 265 313 379 391 400 417 466 478- 479 487 496 519 521 523 561 570 583 590 669 728 764 784 838 842 910 941 950 965 970
mammary gland	Invitrogen	MMG001	4 19 23 26 29 34-39 43 45 48 55 64 66 74 78 87 96-97 114 116 126 129 136 142 149 151 155-156 160 164 168 173 175-176 178 180-181 183 192 197-200 202 204 207-208 215 222 226-228 230 232 235-238 242 246 248 250 252-257 261-262 268 272 274 278 280 301 303 322 329 335 337 343 363 368-371 374 379 381 391 397 400-401 417 426 429 431 437 439-441 443 445 449- 450 455 464 475 478-479 484-485 487-488 498-499 504 507 512 517 519 523 526 532-533 553 557 565 570-571 573 575 577-578 590-591 606 617 619 636 640 646 648 663 677-678 680 691 697 702 708 711 732 744 764 792 803 811-813 817 875-877 885 887-888 900 902 905 908 910-911 918 921-922 934 937 939 941-942 946 951 958 960 965 968 983 989 993 999 1003 1008
induced neuron cells	Strategene	NTD001	39 122 148 152 181 212 246 266 313 337 358 379 452 467 479 484 519 553 561 583 621-626 680 872 881 910 924 941
retinoid acid induced neuronal cells	Strategene	NTR001	37 148 152 168 541 583
neuronal cells	Strategene	NTU001	29 37 147 202 221-222 237 246 262 337 361 391 400 429 439 460 487 504 526 541 583 772 816 924 945 965
pituitary gland	Clontech	PIT004	391 396 764
placenta	Clontech	PLA003	123 183 544 803
prostate	Clontech	PRT001	60-61 76 96 122 145-148 153-154 175 178 183 201 204 226 228 235 237 241 245 248 250-251 256 262 265 280 284 324-325 337 397 400 409 436-437 456 464 478 480 487 489-490 492 508 516-517 524 552 561 583 605 722 740 747 849 889 906 924 926 939 958 974 1005
rectum	Invitrogen	REC001	26 29 43 48 70 74 80 108 114 135- 136 140 168 178-179 208 226 257

TABLE 1

TISSUE ORIGIN	RNA SOURCE	HYSEQ LIBRARY NAME	SEQ ID NOS: OF NUCLEOTIDE(S)
			262 346 348 371 379 411 413 436-437 475 479 484 499 504 517 526 534 548-549 555 570 577-578 606 636 697 729 764 778 793 885 900 906 908 910 937 941 951 965 989 999
salivary gland	Clontech	SAL001	7 38 43 74 87 98 112 122 136 142 148 162 169 181 183-185 207 215 228 235 250 254-255 265 280 349-350 394 437 443 464 508 515-516 519 559 598 614 619 658 666-667 680 724 762-763 771 803 816 842 930 933-934 953
salivary gland	Clontech	SALS03	48 108 515 617 900
skin fibroblast	ATCC	SFB001	39
skin fibroblast	ATCC	SFB002	222 803
skin fibroblast	ATCC	SFB003	237
small intestine	Clontech	SIN001	16 19 29 39 48 56 65 73 96 108 122 136 148 152 155 160 162 165 168 172 181 191 208 234 244 246 266 282 296 379 394 431 440 443 464 479-480 484 519 571 578 583 617 619 648 662 694 703 752 763 806 838 908 910 926 937 941 966 972 976
skeletal muscle	Clontech	SKM001	34 112 116 147 149 152 163 167 373 379 484 515 553 561-562 781 838 910 941
spinal cord	Clontech	SPC001	19 22 29 31 55 58 70-71 78 122 134 145 148 150 152 159-160 163 166 171 175-176 183 200-201 203-204 220 222 224 235 237 246 248 250 257 262 266-268 279-280 327-328 330 337 343 346 371 379 389 396 416 429-430 437 443 452-453 456 467 475 479 493-494 498 500 502 541 544 553 561 583 619 635-636 638 640 680 682 696 764 785 900 902 910 941 950 982 994
adult spleen	Clontech	SPLC01	254 529 701
stomach	Clontech	STO001	48 53 72 74 122 142 152 161 178 181 200-202 204 208 240 251 254 265 268 309 347 397 410 437 512 539 550 583 616 636 657 659 720 722 921
thalamus	Clontech	THA002	35 53 78 114 123 156 176 181 228 235 246 252 255-256 265 280 329 331 343 379 437 452 457 467 479 484 496 507 519 553 571 593 619 692 723 754 758 764 853 910 925 941 950 967 981 1003
thymus	Clontech	THM001	29 78 112 122 148 151 160-161 169 176 180-181 183 188 198 201 204-206 212 250 254 313 374 379 397 412 429 437 446 453 471-472 484 513 521 529 552-553 561 565 619 636 666 708 739 742 764 771 816

TABLE 1

TISSUE ORIGIN	RNA SOURCE	HYSEQ LIBRARY NAME	SEQ ID NOS: OF NUCLEOTIDE(S)
			838 910 941-942 944 947 958 969 979 982 989 999 1007
thymus	Clontech	THMc02	9 19 32 36 63 67 74 78 80 85-86 122-123 138 142 145 147-148 160- 161 169 175-176 181 183-184 187 194 198 202 204 208 211 238 244 246 250 252-254 257 262 265 270- 271 283-285 317 333 349 359-360 379 400-401 406 413 418 429 431 433 436 440-441 473 479 484 487 512-513 517-518 523 525 529 533 535-537 541 544 553 556 561 565 567-570 572-573 578 583 615-619 636 644 660-661 681 683 687 698 732 739 763-764 783 785 789 807- 808 811 816 842 852 864 868-869 900 904 906 910 924 926 930 938 941 965 968 974 979 992 1006-1007
thyroid gland	Clontech	THR001	5 10 13-14 19 23 35 37 39 47 59- 61 64 74 79 87 100 110 112 117 122-123 133 141-142 145 148 152 156 160 168 181 187 199-202 204- 205 207-208 210 220 224-225 228 234-235 237 246-247 251-252 254- 256 262 265 267-268 280-281 284 286 301 308 325 332-333 335 337 343 346 363 371 374 378-379 383 394 396-397 400 420 429 431-432 436 445 452 456 464 467-468 474 479-480 484 487 492 499 507 519 522 533 537 550 553 559 561 569 583 619 638 650 653 655 672 678 680 692 705 719 727 748 764 766- 767 769 792 797 816 821 854 906 910-911 921 924 926 928 941 946 951 958 960-961 967 971 974-975 978 984 989 999
trachea	Clontech	TRC001	43 48 108 112 142 148 168 204 208 212 221-222 254 265 282 286 317 371 382 425 440 501 553 565 910
uterus	Clontech	UTR001	1 37 39 62 145 148 163 183 188 200 257 265 268 346 372 405 408 420 431 520 538 561-562 571 640 680 711 842 850-851 885 910 957

TABLE 2

SEQ ID NO: OF NUCLEOTIDE	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH- WATERMAN SCORE	% IDENTITY
1	AF208846	Homo sapiens	BM-004	172	43
2	Y53871	Homo sapiens	A human brain-derived signalling factor polypeptide.	574	99
3	AE003620	Drosophila melanogaster	CG8486 gene product	112	33
4	AF1.93807	Homo sapiens	Rh type B glycoprotein	1204	96
5	Y87156	Homo sapiens	Human secreted protein sequence SEQ ID NO:195.	89	46
6	Y71.062	Homo sapiens	Human membrane transport protein, MTRP-7.	135	30
7	AB047936	Macaca fascicularis	hypothetical protein	81	38
8	Y36156	Homo sapiens	Human secreted protein #28.	158	68
9	AB040964	Homo sapiens	KIAA1531 protein	495	100
10	U29725	Homo sapiens	BMK1 alpha kinase	114	35
11	X00822	Gallus gallus	collagen type III	54	52
12	Y27868	Homo sapiens	Human secreted protein encoded by gene No. 107.	119	43
13	W74813	Homo sapiens	Human secreted protein encoded by gene 85 clone HSDFV29.	722	92
14	W74813	Homo sapiens	Human secreted protein encoded by gene 85 clone HSDFV29.	722	92
15	AF119851	Homo sapiens	PRO1722	333	70
16	AF264750	Homo sapiens	ALR-like protein	133	100
17	X91014	Mus musculus	alpha 1 type XI collagen	131	72
18	AF090930	Homo sapiens	PRO0478	109	90
19	Y86456	Homo sapiens	Human gene 46-encoded protein fragment, SEQ ID NO:371.	618	95
20	AF084535	Homo sapiens	laforin	1809	100
21	Y27585	Homo sapiens	Human secreted protein encoded by gene No. 19.	587	98
22	Z68748	Caenorhabditis elegans	Similairity to Yeast hypothetical protein YEH4 (SW:YEH4_YEAST) ~cDNA A EST yk87c11.3 comes from this gene~cDNA EST yk87c11.5 comes	214	37

TABLE 2

SEQ ID NO: OF NUCLEOTIDE	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH- WATERMAN SCORE	% IDENTITY
			from this gene~cDNA EST yk497d5.3 comes from this gene~cDNA EST yk186a5.5 comes from this gene~cDNA EST yk243b10.5 comes from this gene~cDNA EST yk497d5.5 comes from this gene		
23	D86973	Homo sapiens	similar to Yeast translation activator GCN1 (P1:A48126)	12053	100
24	Y09945	Rattus norvegicus	putative integral membrane transport protein	458	50
25	U25739	Mus musculus	YSPL-1 form 1	719	77
26	AK024427	Homo sapiens	FLJ00016 protein	668	100
27	AP001707	Homo sapiens	human gene for claudin-8, Accession No. AJ250711	603	100
28	U16030	Brugia malayi	cuticular collagen Bmcoll-2	78	37
29	G02479	Homo sapiens	Human secreted protein, SEQ ID NO: 6560.	442	100
30	Y13375	Homo sapiens	Amino acid sequence of protein PRO262.	1806	99
31	AF077226	Homo sapiens	copine III	1757	65
32	W75198	Homo sapiens	Human secreted protein encoded by gene 3 clone HCEDO84.	208	100
33	AF151978	Homo sapiens	amino acid transporter B0+	3436	100
34	Y66735	Homo sapiens	Membrane-bound protein PRO1153.	1006	100
35	AC003093	Homo sapiens	OXYSTEROOL-BINDING PROTEIN; 45% similarity to P22059 (PID:gl29308)	764	60
36	AF286861	Fasciola hepatica	tegumental antigen-like protein	79	30
37	AF201945	Homo sapiens	HNOEL-iso	2152	100
38	AF258465	Homo sapiens	OTRPC4	1668	99
39	AF173003	Homo sapiens	apoptosis regulator	2421	100
40	Y53023	Homo sapiens	Human secreted protein clone qf662_3 protein sequence SEQ ID	128	41

TABLE 2

SEQ ID NO: OF NUCLEOTIDE	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH- WATERMAN SCORE	% IDENTITY
			NO:52.		
41	M25750	Oryctolagus cuniculus	sarcolumenin precursor	2307	97
42	G03797	Homo sapiens	Human secreted protein, SEQ ID NO: 7878.	186	75
43	X57805	Homo sapiens	immunoglobulin lambda light chain	1102	91
44	AE003689	Drosophila melanogaster	CG4596 gene product	419	44
45	Y50934	Homo sapiens	Human fetal brain cDNA clone vc30_1 derived protein #1.	644	100
46	Y19562	Homo sapiens	Amino acid sequence of a human secreted protein.	80	45
47	AF016272	Homo sapiens	Ksp-cadherin	4263	99
48	R13111	Homo sapiens	1B1 IgG aberrant light chain with duplicated variable region.	1000	92
49	AK001636	Homo sapiens	unnamed protein product	1630	97
50	Y65155	Homo sapiens	Human 5' EST related polypeptide SEQ ID NO:1316.	78	34
51	G00471	Homo sapiens	Human secreted protein, SEQ ID NO: 4552.	281	91
52	AJ272050	Homo sapiens	transcription initiation factor IA protein	165	68
53	Y42388	Homo sapiens	Amino acid sequence of pt127_1.	668	73
54	AF193807	Homo sapiens	Rh type B glycoprotein	248	97
55	AF132611	Homo sapiens	monocarboxylate transporter MCT3	139	37
56	U43940	Rattus norvegicus	focal adhesion kinase	141	84
57	L17318	Rattus norvegicus	proline-rich proteoglycan	124	37
58	G02832	Homo sapiens	Human secreted protein, SEQ ID NO: 6913.	132	48
59	G00357	Homo sapiens	Human secreted protein, SEQ ID NO: 4438.	95	64
60	Y12723	Homo sapiens	Human 5' EST secreted protein SEQ ID NO:313.	91	50
61	Y19450	Homo sapiens	Amino acid sequence of a human secreted	406	100

TABLE 2

SEQ ID NO: OF NUCLEOTIDE	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH- WATERMAN SCORE	% IDENTITY
			protein.		
62	AF156549	Mus musculus	putative E1-E2 ATPase	876	65
63	AL356276	Homo sapiens	ba367J7.5 (novel Immunoglobulin domain containing protein)	655	84
64	AL133105	Homo sapiens	hypothetical protein	1783	99
65	U32189	Oryctolagus cuniculus	histidine-rich glycoprotein precursor	73	40
66	Y91433	Homo sapiens	Human secreted protein sequence encoded by gene 33 SEQ ID NO:154.	758	98
67	W75198	Homo sapiens	Human secreted protein encoded by gene 3 clone HCEDO84.	208	100
68	AF020651	Homo sapiens	T cell receptor alpha chain variable region	742	93
69	AF118086	Homo sapiens	PRO1992	158	61
70	X52454	Drosophila melanogaster	rho	224	36
71	W40353	Homo sapiens	Human unspecified protein from US5702907.	146	67
72	Y66690	Homo sapiens	Membrane-bound protein PRO813.	971	98
73	AJ002744	Homo sapiens	UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferase 7	1518	98
74	AC024792	Caenorhabditis elegans	contains similarity to TR:P78316	423	36
75	AB016088	Homo sapiens	RNA binding protein	109	32
76	Y94953	Homo sapiens	Human secreted protein clone fy356_14 protein sequence SEQ ID NO:112.	2484	100
77	AF107406	Homo sapiens	GW128	74	51
78	Y13401	Homo sapiens	Amino acid sequence of protein PRO339.	1681	96
79	Y94290	Homo sapiens	Human myosin heavy chain homologue.	1819	99
80	AF007194	Homo sapiens	mucin	4875	100
81	AF229179	Homo sapiens	kidney-specific membrane protein NX-17	949	99

TABLE 2

SEQ ID NO: OF NUCLEOTIDE	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH- WATERMAN SCORE	% IDENTITY
82	AL356173	Neurospora crassa	hypothetical protein	83	29
83	G00437	Homo sapiens	Human secreted protein, SEQ ID NO: 4518.	87	69
84	K03036	Mus musculus	alpha-1 type I procollagen	114	38
85	AF233261	Homo sapiens	otoraplin	676	100
86	AF073519	Homo sapiens	small EDRK-rich factor 1, long isoform	100	45
87	AC021640	Arabidopsis thaliana	putative phosphatidate phosphohydrolase	387	43
88	AB040812	Homo sapiens	protein kinase PAK5	1159	100
89	AL365409	Homo sapiens	similar to (NP_034322.1 ) sex- determination protein homolog Fem1a	694	100
90	U81035	Rattus norvegicus	ankyrin binding cell adhesion molecule neurofascin	189	63
91	W88684	Homo sapiens	Secreted protein encoded by gene 151 clone HNHED86.	134	65
92	Y66734	Homo sapiens	Membrane-bound protein PRO1097.	297	70
93	AB031051	Homo sapiens	organic anion transporter OATP-E	283	40
94	B08976	Homo sapiens	Human secreted protein sequence encoded by gene 28 SEQ ID NO:133.	71	27
95	U83115	Homo sapiens	non-lens beta gamma-crystallin like protein	245	97
96	AF156551	Mus musculus	putative E1-E2 ATPase	3779	86
97	AF062476	Mus musculus	retinoic acid- responsive protein; STRA6	1091	74
98	Y87072	Homo sapiens	Human secreted protein sequence SEQ ID NO:111.	490	100
99	AF116652	Homo sapiens	PRO0813	1015	99
100	AF159567	Homo sapiens	C2H2 (Kruppel-type) zinc finger protein	2176	100
101	D25328	Homo sapiens	platelet-type phosphofructokinase	109	95
102	AB018563	Homo sapiens	TML1	98	68
103	X83107	Homo sapiens	bmx	232	85

TABLE 2

SEQ ID NO: OF NUCLEOTIDE	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH- WATERMAN SCORE	% IDENTITY
104	U49973	Homo sapiens	ORF1; MER37; putative transposase similar to pogo element	131	43
105	Y86472	Homo sapiens	Human gene 52- encoded protein fragment, SEQ ID NO:387.	150	54
106	AF020276	Homo sapiens	spinocerebellar ataxia 7	96	37
107	W57901	Homo sapiens	Protein of clone CT748_2.	1499	96
108	R13111	Homo sapiens	1B1 IgG aberrant light chain with duplicated variable region.	1210	84
109	W50192	Homo sapiens	Amino acid sequence of salivary protein CON-1.	95	32
110	AB046634	Macaca fascicularis	hypothetical protein	282	75
111	AF242432	Mus musculus	neuronal apoptosis inhibitory protein 6	486	29
112	AB000280	Rattus norvegicus	peptide/histidine transporter	2490	88
113	AF182443	Rattus norvegicus	F-box protein FBL2	597	99
114	AJ245874	Homo sapiens	putative ATG/GTP binding protein	1242	100
115	AF179828	Saimiri sciureus	olfactory receptor	444	66
116	Y66735	Homo sapiens	Membrane-bound protein PRO1153.	1006	100
117	Y94344	Homo sapiens	Human cell surface receptor protein #11.	892	90
118	AJ238706	Drosophila melanogaster	monocarboxylate transporter 1 homologue	226	31
119	AF180728	Drosophila melanogaster	sulfate transporter	312	45
120	AE004890	Pseudomonas aeruginosa	L-lactate permease	534	89
121	X91837	Saccharomyces cerevisiae	cell division cycle protein CDC55	435	98
122	U93565	Homo sapiens	putative p150	1911	90
123	AJ000332	Homo sapiens	Glucosidase II	5043	99
124	AF204674	Homo sapiens	muscle disease-related protein	377	72
125	S58722	Homo sapiens	X-linked retinopathy protein {C-terminal, clone}	196	68

TABLE 2

SEQ ID NO: OF NUCLEOTIDE	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH- WATERMAN SCORE	% IDENTITY
			XEH.8c}		
126	S58722	Homo sapiens	X-linked retinopathy protein {C-terminal, clone XEH.8c}	196	68
127	J03848	Mesocricetus auratus	metallothionein II	147	51
128	G02994	Homo sapiens	Human secreted protein, SEQ ID NO: 7075.	93	64
129	AF116238	Homo sapiens	pseudouridine synthase 1	1927	99
130	G03411	Homo sapiens	Human secreted protein, SEQ ID NO: 7492.	183	65
131	AF222861	Sus scrofa	type X collagen	90	34
132	G03628	Homo sapiens	Human secreted protein, SEQ ID NO: 7709.	60	66
133	Y10529	Homo sapiens	olfactory receptor	766	61
134	AF164612	Homo sapiens	Gag protein	125	43
135	Y12713	Mus musculus	Pro-Pol-dUTPase polyprotein	181	47
136	X57816	Homo sapiens	immunoglobulin lambda light chain	550	57
137	U07808	Mus musculus	metallothionein IV	55	37
138	AB031227	Pisum sativum	PsAD1	68	50
139	AB035520	Oryctolagus cuniculus	parchorin	1324	57
140	AB007891	Homo sapiens	KIAA0431	117	46
141	Y00278	Homo sapiens	Human secreted protein encoded by gene 21.	234	92
142	Y68810	Homo sapiens	A rat heavy chain region and a human hinge region.	1124	92
143	M58526	Homo sapiens	alpha-5 type IV collagen	4597	97
144	AF119851	Homo sapiens	PRO1722	192	66
145	X84908	Homo sapiens	phosphorylase kinase	3798	97
146	Y76155	Homo sapiens	Human secreted protein encoded by gene 32.	81	52
147	U13766	Murine leukemia virus	gag-pol polyprotein	735	36
148	AF034198	Homo sapiens	IGSF1	7154	100
149	Y94343	Homo sapiens	Human cell surface receptor protein #10.	1331	100
150	Y87211	Homo sapiens	Human secreted	759	97

TABLE 2

SEQ ID NO: OF NUCLEOTIDE	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH- WATERMAN SCORE	% IDENTITY
			protein sequence SEQ ID NO: 250.		
151	AJ252258	human herpesvirus 2	glycoprotein G-2	115	30
152	V00662	Homo sapiens	URF 1 (NADH dehydrogenase subunit)	1283	85
153	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	142	61
154	A23786	Beta vulgaris	chitinase 1	138	41
155	Z34465	Zea mays	extensin-like protein	97	36
156	X79389	Homo sapiens	glutathione transferase T1	721	66
157	M22333	Homo sapiens	unknown protein	106	46
158	AL118502	Homo sapiens	bA371L19.1 (novel protein)	2471	100
159	AJ012582	Homo sapiens	hyperpolarization- activated cation channel HCN2	3076	100
160	D26351	Homo sapiens	human type 3 inositol 1,4,5- trisphosphate receptor	8901	99
161	AF067656	Homo sapiens	ZW10 interactor Zwint	951	97
162	AE003461	Drosophila melanogaster	CG11300 gene product	76	29
163	Y48518	Homo sapiens	Human breast tumour-associated protein 63.	355	100
164	G00517	Homo sapiens	Human secreted protein, SEQ ID NO: 4598.	83	34
165	G03786	Homo sapiens	Human secreted protein, SEQ ID NO: 7867.	251	53
166	Y00765	Homo sapiens	Prion protein CJAS.	63	37
167	Y21050	Homo sapiens	Human glial fibrillary acidic protein GFAP mutant fragment 59.	206	71
168	X74929	Homo sapiens	Keratin 8	1462	95
169	U29488	Caenorhabditis elegans	similar to DNAJ protein	555	29
170	L27428	Homo sapiens	reverse transcriptase	145	45
171	W19932	Homo sapiens	Alzheimer's disease protein encoded by DNA from plasmid pGCS55.	362	100
172	AF178983	Homo sapiens	Ras-associated	497	100

TABLE 2

SEQ ID NO: OF NUCLEOTIDE	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH- WATERMAN SCORE	% IDENTITY
			protein Rap1		
173	U70136	Homo sapiens	megakaryocyte stimulating factor; MSF	206	28
174	G00352	Homo sapiens	Human secreted protein, SEQ ID NO: 4433.	109	64
175	U28143	Gallus gallus	synemin	1014	39
176	Y13401	Homo sapiens	Amino acid sequence of protein PRO339.	1978	96
177	AJ243396	Homo sapiens	voltage-gated sodium channel beta-3 subunit	947	99
178	M77812	Oryctolagus cuniculus	myosin heavy chain	4079	98
179	AF200344	Homo sapiens	aspartyl protease 3	956	91
180	AF200815	Homo sapiens	FUSED serine/threonine kinase	1597	99
181	G03786	Homo sapiens	Human secreted protein, SEQ ID NO: 7867.	147	83
182	Y00313	Homo sapiens	Human secreted protein encoded by gene 56.	56	29
183	X00699	Homo sapiens	precursor	583	66
184	AF269289	Homo sapiens	unknown	81	32
185	G03797	Homo sapiens	Human secreted protein, SEQ ID NO: 7878.	176	66
186	Y20298	Homo sapiens	Human apolipoprotein E mutant protein fragment 11.	110	34
187	AF161437	Homo sapiens	HSPC319	867	99
188	Y19684	Homo sapiens	SEQ ID NO 402 from WO9922243.	124	47
189	Y74050	Homo sapiens	Human prostate tumor EST fragment derived protein #237.	78	42
190	Y08986	Brassica napus	oleosin-like protein	106	36
191	AF119851	Homo sapiens	PRO1722	173	66
192	AF116712	Homo sapiens	PRO2738	166	50
193	AF186084	Homo sapiens	epidermal growth factor repeat containing protein	2022	85
194	M59819	Homo sapiens	granulocyte colony-stimulating factor receptor	4232	100
195	Y86228	Homo sapiens	Human secreted protein HFXJX44,	250	100

TABLE 2

SEQ ID NO: OF NUCLEOTIDE	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH- WATERMAN SCORE	% IDENTITY
			SEQ ID NO:143.		
196	Y45382	Homo sapiens	Human secreted protein fragment encoded from gene 28.	181	63
197	X94991	Homo sapiens	zyxin	566	41
198	M17236	Homo sapiens	MHC HLA-DQ alpha precursor	896	84
199	AC004659	Homo sapiens	BC62940_2	805	53
200	X14420	Homo sapiens	prepro-alpha-1 type 3 collagen	5521	99
201	AF180473	Homo sapiens	Not2p	1628	98
202	X85237	Homo sapiens	human splicing factor	1145	100
203	AL390114	Leishmania major	extremely cysteine/valine rich protein	309	58
204	D42138	Homo sapiens	PIG-B	1479	98
205	Y00062	Homo sapiens	precursor polypeptide (AA -23 to 1120)	3334	98
206	W93946	Homo sapiens	Human regulatory molecule HRM-2 protein.	1011	100
207	AB017563	Homo sapiens	IGSF4	2062	99
208	X54637	Homo sapiens	protein tyrosine kinase	5694	98
209	AF255910	Homo sapiens	vascular endothelial junction-associated molecule	1508	98
210	AF061324	Homo sapiens	sulfonylurea receptor 2A	7545	97
211	U93568	Homo sapiens	p40	197	50
212	AF250842	Drosophila melanogaster	multiple asters	506	32
213	X81479	Homo sapiens	EMR1	4469	99
214	X77748	Homo sapiens	metabotropic glutamate receptor type 3 (mGluR3)	4471	99
215	M60396	Homo sapiens	transcobalamin II	2218	99
216	W48351	Homo sapiens	Human breast cancer related protein BCRB2.	170	71
217	Y36203	Homo sapiens	Human secreted protein #75.	156	73
218	AF119851	Homo sapiens	PRO1722	144	63
219	AJ246002	Mus musculus	spastin protein orthologue	143	100
220	D49958	Homo sapiens	membrane glycoprotein M6	616	57
221	X83573	Homo sapiens	ARSE	2114	93

TABLE 2

SEQ ID NO: OF NUCLEOTIDE	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH- WATERMAN SCORE	% IDENTITY
222	AF126062	Homo sapiens	Arf-like 2 binding protein BART1	508	84
223	L22695	Canine oral papillomavirus	5' end derived by splicing; putative	83	51
224	R95913	Homo sapiens	Neural thread protein.	262	64
225	AP001306	Arabidopsis thaliana	contains similarity to cell wall-plasma membrane linker protein~gene_id:MKA 23.3	79	34
226	G01984	Homo sapiens	Human secreted protein, SEQ ID NO: 6065.	252	64
227	X04614	human herpesvirus 1	IE110	83	35
228	AF151877	Homo sapiens	CGI-119 protein	1203	94
229	AF181467	Homo sapiens	protein Z-dependent protease inhibitor precursor	1483	88
230	Z81326	Homo sapiens	neuroserpin	1763	99
231	AF111173	Homo sapiens	sodium/hydrogen exchanger isoform 5	3512	99
232	X67055	Homo sapiens	inter-alpha-trypsin inhibitor heavy chain H3	4429	98
233	AB004064	Homo sapiens	tomoregulin	1783	98
234	AL096772	Homo sapiens	dJ365O12.1 (KIAA0758 protein)	5465	98
235	X83378	Homo sapiens	putative chloride channel	1620	99
236	AF043644	Homo sapiens	receptor protein tyrosine phosphatase	5127	97
237	AF208536	Homo sapiens	nucleotide binding protein; NBP	1372	100
238	AC005625	Homo sapiens	R27328_1	2435	93
239	X55687	Lycopersicon esculentum	extensin (class II)	58	50
240	M23315	Sesbania rostrata	nodulin	61	36
241	AF102851	Homo sapiens	dolichyl-P-Glc:Man9GlcNAc2-PP-dolichyl glucosyltransferase	1881	99
242	G03793	Homo sapiens	Human secreted protein, SEQ ID NO: 7874.	202	67
243	G03258	Homo sapiens	Human secreted protein, SEQ ID NO: 7339.	203	69
244	AF048774	Homo sapiens	anti-HER3 scFv	903	81

TABLE 2

SEQ ID NO: OF NUCLEOTIDE	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH- WATERMAN SCORE	% IDENTITY
245	AF102851	Homo sapiens	dolichyl-P-Glc:Man9GlcNAc2-PP-dolichyl glucosyltransferase	1867	98
246	L00352	Homo sapiens	low density lipoprotein receptor	3980	100
247	Y79510	Homo sapiens	Human carbohydrate-associated protein CRBAP-6.	1394	100
248	AF202636	Homo sapiens	angiopoietin-like protein PP1158	2164	100
249	X66533	Homo sapiens	guanylate cyclase	1641	97
250	M20504	Homo sapiens	MHC HLA-DR-beta-2 precursor	750	70
251	AF157326	Homo sapiens	TIP120 protein	4278	99
252	M25865	Homo sapiens	von Willebrand factor	10841	95
253	AC005625	Homo sapiens	R27328_1	2435	93
254	A21385	synthetic construct	heavy chain antibody 3D6	1786	94
255	AF182414	Homo sapiens	MDS013	310	48
256	Y54041	Homo sapiens	Protein encoded by a gene reduced in metastatic melanoma cells (grmm-1).	1267	84
257	AJ011415	Homo sapiens	plexin-B1/SEP receptor	1580	60
258	W55030	Homo sapiens	G-protein coupled receptor, long form.	1493	100
259	AF227747	Homo sapiens	voltage-dependent calcium channel alpha 1G subunit isoform bc	6158	100
260	AF111173	Homo sapiens	sodium/hydrogen exchanger isoform 5	3512	99
261	G01984	Homo sapiens	Human secreted protein, SEQ ID NO: 6065.	175	70
262	Y00815	Homo sapiens	put. LAR preprotein (AA -16 to 1881)	5648	100
263	Z34979	Homo sapiens	Human FIZZ3 (inhibitor of neurotrophin action) cDNA.	582	100
264	AF119851	Homo sapiens	PRO1722	189	73
265	AL049798	Homo sapiens	dJ797M17.1 (Dermatopontin)	1007	99
266	AL035684	Homo sapiens	dJ1114A1.1 (KIAA0611 (putative E1-E2 ATPase) protein)	1978	99

TABLE 2

SEQ ID NO: OF NUCLEOTIDE	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH- WATERMAN SCORE	% IDENTITY
267	U49055	Rattus norvegicus	rA8	4382	87
268	X15332	Homo sapiens	alpha-1 (III) collagen	4170	99
269	Z98884	Homo sapiens	dJ467L1.1 (KIAA0833)	2010	100
270	AF085244	Homo sapiens	C2H2 type Kruppel- like zinc finger protein splice variant b	7331	98
271	Y00319	Homo sapiens	Human secreted protein encoded by gene 63.	214	82
272	X04434	Homo sapiens	IGF-I receptor	5832	99
273	AC005626	Homo sapiens	R29124_1	1129	89
274	X52046	Mus musculus	type III collagen	819	37
275	M22207	Tripneustes gratilla	217g protein	168	51
276	M32317	Homo sapiens	HLA protein allele B7	1536	84
277	L05485	Homo sapiens	surfactant protein D	1693	87
278	W88504	Homo sapiens	Human epidermoid carcinoma clone HP10428-encoded membrane protein.	1187	100
279	AF078850	Homo sapiens	steroid dehydrogenase homolog	794	100
280	X83378	Homo sapiens	putative chloride channel	1620	99
281	AL035701	Homo sapiens	dJ8B1.3 (similar to PLASMA-CELL MEMBRANE GLYCOPROTEIN PC-1)	2412	99
282	Y87068	Homo sapiens	Human secreted protein sequence SEQ ID NO:107.	528	100
283	L40806	Neurospora crassa	Restriction enzyme inactivation of met-10 complementation in this region. Sequence similarity to <i>S. cerevisiae</i> chromosome VIII cosmid 9205, accession no. U10556 CDS residues 22627-24126	536	35
284	W88552	Homo sapiens	Secreted protein encoded by gene 19 clone HSAVU34.	3078	99

TABLE 2

SEQ ID NO: OF NUCLEOTIDE	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH- WATERMAN SCORE	% IDENTITY
285	G03790	Homo sapiens	Human secreted protein, SEQ ID NO: 7871.	108	50
286	X68060	Homo sapiens	DNA topoisomerase II	8296	99
287	G00352	Homo sapiens	Human secreted protein, SEQ ID NO: 4433.	114	41
288	AC004602	Homo sapiens	F23487_2	202	49
289	AF196329	Homo sapiens	triggering receptor expressed on monocytes 1	1211	99
290	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	202	62
291	G03043	Homo sapiens	Human secreted protein, SEQ ID NO: 7124.	93	62
292	Y12550	Homo sapiens	Human 5' EST secreted protein SEQ ID NO: 215 from WO 9906553.	141	100
293	D43756	Canis familiaris	fibrinogen A-alpha-chain	102	33
294	U38545	Homo sapiens	phospholipase D1	5681	99
295	W42076	Homo sapiens	The amino acid sequence of the O276_16 protein.	236	100
296	AF090930	Homo sapiens	PRO0478	128	60
297	Y64747	Homo sapiens	Human 5' EST related polypeptide SEQ ID NO:908.	471	98
298	G01234	Homo sapiens	Human secreted protein, SEQ ID NO: 5315.	280	71
299	G02514	Homo sapiens	Human secreted protein, SEQ ID NO: 6595.	94	76
300	G02493	Homo sapiens	Human secreted protein, SEQ ID NO: 6574.	112	46
301	Z38061	Saccharomyces cerevisiae	mal5, stal, len: 1367, CAI: 0.3, AMYH_YEAST P08640 GLUCOAMYLASE S1 (EC 3.2.1.3)	340	27
302	Y59672	Homo sapiens	Secreted protein 108-006-5-0-E6-FL.	530	78
303	Y95018	Homo sapiens	Human secreted protein vp19_1, SEQ ID NO:76.	76	35
304	W34623	Homo sapiens	Human C3 protein mutant FT-1.	117	46

TABLE 2

SEQ ID NO: OF NUCLEOTIDE	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH- WATERMAN SCORE	% IDENTITY
305	Y87292	Homo sapiens	Human signal peptide containing protein HSPP-69 SEQ ID NO:69.	81	50
306	AF210651	Homo sapiens	NAG18	135	60
307	Y14482	Homo sapiens	Fragment of human secreted protein encoded by gene 17.	212	58
308	Y76325	Homo sapiens	Fragment of human secreted protein encoded by gene 35.	343	93
309	Y36156	Homo sapiens	Human secreted protein #28.	203	75
310	AF090931	Homo sapiens	PRO0483	76	50
311	AC004943	Homo sapiens	alpha-fetoprotein enhancer-binding protein; 99% identical to A41948 (PID:g283975)	351	85
312	G02558	Homo sapiens	Human secreted protein, SEQ ID NO: 6639.	144	52
313	AK000128	Homo sapiens	unnamed protein product	1338	100
314	G03786	Homo sapiens	Human secreted protein, SEQ ID NO: 7867.	164	83
315	AF090942	Homo sapiens	PRO0657	253	68
316	AF116712	Homo sapiens	PRO2738	181	52
317	AF043726	Mus musculus	PHD-finger protein	1605	64
318	Y99368	Homo sapiens	Human PRO1326 (UNQ686) amino acid sequence SEQ ID NO:100.	145	51
319	AF065314	Homo sapiens	cone photoreceptor cGMP-gated channel alpha subunit	292	98
320	AF003389	Caenorhabditis elegans	contains similarity to N-chimaerins	162	28
321	Y66755	Homo sapiens	Membrane-bound protein PRO1185.	993	100
322	AF109906	Mus musculus	RD	118	69
323	AF199323	Rattus norvegicus	RIM2-2A	364	85
324	G02538	Homo sapiens	Human secreted protein, SEQ ID NO: 6619.	104	65
325	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	138	65
326	Y41266	Homo sapiens	Human T139 protein.	591	100
327	G02920	Homo sapiens	Human secreted protein, SEQ ID NO:	103	67

TABLE 2

SEQ ID NO: OF NUCLEOTIDE	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH- WATERMAN SCORE	% IDENTITY
			7001.		
328	G00636	Homo sapiens	Human secreted protein, SEQ ID NO: 4717.	80	36
329	U37769	Oryctolagus cuniculus	protein phosphatase 2A0 B' regulatory subunit alpha isoform	556	88
330	AE001424	Plasmodium falciparum	RESA-H3 antigen	208	21
331	AF090930	Homo sapiens	PRO0478	156	82
332	AF161356	Homo sapiens	HSPC093	169	64
333	G04055	Homo sapiens	Human secreted protein, SEQ ID NO: 8136.	425	100
334	D79985	Homo sapiens	putative hydrophobic domain in the central region.	371	86
335	Y41401	Homo sapiens	Human secreted protein encoded by gene 94 clone HLYCH68.	392	100
336	W18651	Homo sapiens	Human apolipoprotein E gene +1 frameshift mutant product.	478	88
337	Y20921	Homo sapiens	Human presenilin II wild type protein fragment 5.	2126	96
338	AF010144	Homo sapiens	neuronal thread protein AD7c-NTP	233	75
339	D28500	Homo sapiens	mitochondrial isoleucine tRNA synthetase	175	89
340	Y13357	Homo sapiens	Amino acid sequence of protein PRO227.	148	50
341	AL096677	Homo sapiens	dJ322G13.2 (similar to cystatin)	94	50
342	Y10843	Homo sapiens	Amino acid sequence of a human secreted protein.	186	86
343	X54134	Homo sapiens	protein-tyrosine phosphatase	3705	100
344	Z33908	Mus musculus	inositol 1,4,5-trisphosphate receptor	315	84
345	G00241	Homo sapiens	Human secreted protein, SEQ ID NO: 4322.	130	46
346	AF071172	Homo sapiens	HERC2	23705	99
347	AB015346	Homo sapiens	Eps15R	209	95
348	Y48596	Homo sapiens	Human breast	108	34

TABLE 2

SEQ ID NO: OF NUCLEOTIDE	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH- WATERMAN SCORE	% IDENTITY
			tumour-associated protein 57.		
349	G03058	Homo sapiens	Human secreted protein, SEQ ID NO: 7139.	85	66
350	Y73443	Homo sapiens	Human secreted protein clone yb187_1 protein sequence SEQ ID NO:108.	90	36
351	G03793	Homo sapiens	Human secreted protein, SEQ ID NO: 7874.	126	66
352	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	324	73
353	Y64747	Homo sapiens	Human 5' EST related polypeptide SEQ ID NO:908.	527	98
354	AF255342	Homo sapiens	putative pheromone receptor V1RL1 long form	147	59
355	W48351	Homo sapiens	Human breast cancer related protein BCRB2.	85	61
356	G03060	Homo sapiens	Human secreted protein, SEQ ID NO: 7141.	191	72
357	AF124729	Mus musculus	acinusS'	124	31
358	U37352	Homo sapiens	protein phosphatase 2A B'alpha1 regulatory subunit	1016	95
359	AF280605	Triticum aestivum	omega gliadin storage protein	125	35
360	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	150	81
361	AL035398	Homo sapiens	dJ796I17.2 (CGT-51)	226	64
362	AK000307	Homo sapiens	unnamed protein product	882	97
363	Y41401	Homo sapiens	Human secreted protein encoded by gene 94 clone HLYCH68.	392	100
364	AF288480	Homo sapiens	tubby super-family protein	238	87
365	AL023706	Schizosaccharomyces pombe	possible pre-mRNA processing by similarity to yeast prp39	383	34
366	W48351	Homo sapiens	Human breast cancer related protein BCRB2.	85	61

TABLE 2

SEQ ID NO: OF NUCLEOTIDE	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH- WATERMAN SCORE	% IDENTITY
367	S68978	Oryctolagus cuniculus	interleukin-1 receptor antagonist intracellular form	53	58
368	AF047602	Equus zebra hartmannae	luteinizing hormone/chorionic gonadotrophin beta-subunit	68	37
369	AF119851	Homo sapiens	PRO1722	180	75
370	U15195	Homo sapiens	alpha-1 type II collagen	59	43
371	U02082	Homo sapiens	guanine nucleotide regulatory protein	2648	100
372	AF096895	Homo sapiens	chemokine-like factor 1	508	100
373	G03786	Homo sapiens	Human secreted protein, SEQ ID NO: 7867.	315	65
374	AF010144	Homo sapiens	neuronal thread protein AD7c-NTP	240	67
375	U22376	Homo sapiens	alternatively spliced product using exon 13A	191	80
376	U08310	Saimiri sciureus	prion protein	245	66
377	A76867	unidentified	Chimere G.CSF-Gly4-SA en aval region prepro de SAH	550	99
378	G00442	Homo sapiens	Human secreted protein, SEQ ID NO: 4523.	94	53
379	AF010144	Homo sapiens	neuronal thread protein AD7c-NTP	355	53
380	AB023634	Rattus norvegicus	Ca/calmodulin-dependent protein kinase phosphatase	161	91
381	Y99437	Homo sapiens	Human PRO1508 (UNQ761) amino acid sequence SEQ ID NO:336.	805	100
382	W48351	Homo sapiens	Human breast cancer related protein BCRB2.	139	61
383	M58511	Homo sapiens	iron-responsive element-binding protein/iron regulatory protein 2	286	100
384	Y02671	Homo sapiens	Human secreted protein encoded by gene 22 clone HMSJW18.	99	71
385	AJ012166	Canis familiaris	brain-specific synapse associated	86	38

TABLE 2

SEQ ID NO: OF NUCLEOTIDE	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH- WATERMAN SCORE	% IDENTITY
			protein, Bassoon		
386	L07809	Homo sapiens	dynamin	98	31
387	M15530	Homo sapiens	B-cell growth factor	158	69
388	AF090172	Mycoplasma pneumoniae	revertant adhesin-related protein P30	109	31
389	AJ278964	Homo sapiens	cytosolic beta-glucosidase	165	52
390	AF190642	Homo sapiens	phosphoinositide-specific phospholipase C PLC-epsilon	1095	98
391	X13238	Homo sapiens	cytochrome c oxidase subunit VIc preprotein	379	100
392	AF225417	Homo sapiens	88.8 kDa protein	1634	98
393	Y02693	Homo sapiens	Human secreted protein encoded by gene 44 clone HTDAD22.	278	75
394	AF151037	Homo sapiens	HSPC203	554	100
395	AJ276396	Homo sapiens	matrix extracellular phosphoglycoprotein	465	100
396	X51405	Homo sapiens	pre-pro polypeptide (AA -25 to 451)	2536	100
397	W78128	Homo sapiens	Human secreted protein encoded by gene 3 clone HOSBI96.	564	71
398	Y87346	Homo sapiens	Human signal peptide containing protein HSPP-123 SEQ ID NO:123.	290	90
399	G03564	Homo sapiens	Human secreted protein, SEQ ID NO: 7645.	72	52
400	U89436	Homo sapiens	tyrosyl-tRNA synthetase	2719	100
401	W80993	Homo sapiens	Human RIP-interacting factor RIF.	1724	100
402	Y27907	Homo sapiens	Human secreted protein encoded by gene No. 119.	95	59
403	AB033102	Homo sapiens	KIAA1276 protein	921	100
404	G03797	Homo sapiens	Human secreted protein, SEQ ID NO: 7878.	192	55
405	AF096895	Homo sapiens	chemokine-like factor 1	508	100
406	Y29861	Homo sapiens	Human secreted protein clone	791	98

TABLE 2

SEQ ID NO: OF NUCLEOTIDE	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH- WATERMAN SCORE	% IDENTITY
			cb98_4.		
407	Y00293	Homo sapiens	Human secreted protein encoded by gene 36.	237	97
408	W40215	Homo sapiens	Human macrophage antigen.	1358	99
409	J.36056	Homo sapiens	4E-binding protein 2	639	100
410	AJ130710	Homo sapiens	QA79 membrane protein, allelic variant airm-1b	2473	100
411	AF116661	Homo sapiens	PRO1438	146	57
412	W88761	Homo sapiens	Polypeptide fragment encoded by gene 19.	150	58
413	AK024434	Homo sapiens	FLJ00024 protein	574	97
414	Y10376	Homo sapiens	SIRP-beta1	2069	99
415	Y07930	Homo sapiens	Human secreted protein fragment encoded from gene 79.	351	98
416	R99390	Homo sapiens	Human 030 gene (fohy030) product.	804	71
417	AB018253	Rattus norvegicus	voltage-gated ca channel	2419	88
418	AC006017	Homo sapiens	similar to ALR; similar to AAC51735 (PID:g2358287)	2150	97
419	X72925	Homo sapiens	Dsc1b precursor	4390	99
420	AF205940	Homo sapiens	endomucin	1289	100
421	Y27868	Homo sapiens	Human secreted protein encoded by gene No. 107.	134	54
422	W74722	Homo sapiens	Human secreted protein er80_1.	2422	100
423	AF080470	Homo sapiens	pallid	872	100
424	G04072	Homo sapiens	Human secreted protein, SEQ ID NO: 8153.	201	63
425	W90961	Homo sapiens	Human CSGP-1 protein.	869	86
426	M13180	Human herpesvirus 4	nuclear antigen (EBNA 1)	59	45
427	G00365	Homo sapiens	Human secreted protein, SEQ ID NO: 4446.	99	75
428	AF155819	Mus musculus	doublecortin-like kinase	3448	96
429	Y04315	Homo sapiens	Human secreted protein encoded by gene 23.	385	100
430	AB026891	Homo sapiens	cystine/glutamate transporter	2552	100

TABLE 2

SEQ ID NO: OF NUCLEOTIDE	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH- WATERMAN SCORE	% IDENTITY
431	Y15286	Homo sapiens	vacuolar proton-ATPase subunit M9.2	459	100
432	X81053	Homo sapiens	type IV collagen alpha 4 chain	9706	99
433	U41829	Macaca mulatta	MHC class I antigen Mamu B*07	365	76
434	G03371	Homo sapiens	Human secreted protein, SEQ ID NO: 7452.	100	41
435	AF233238	Gallus gallus	BMP signal transducer Smad1	170	74
436	X52425	Homo sapiens	interleukin 4 receptor	4492	99
437	Y06115	Homo sapiens	Human organic cation transporter OCT-3.	2593	96
438	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	130	54
439	L08239	Homo sapiens	located at OATL1	1304	95
440	X17115	Homo sapiens	precursor (AA -15 to 612)	2613	86
441	Y06816	Homo sapiens	Human Notch2 (humN2) protein sequence.	1471	98
442	AB019440	Homo sapiens	immunoglobulin heavy chain variable region	545	88
443	Y87350	Homo sapiens	Human signal peptide containing protein HSPP-127 SEQ ID NO:127.	1061	100
444	AJ271736	Homo sapiens	synaptobrevin-like 1 protein	1128	100
445	Y11534	Homo sapiens	PEG1/MEST	1787	100
446	W85719	Homo sapiens	Novel protein (Clone AJ143_1).	271	100
447	Y07900	Homo sapiens	Human secreted protein fragment encoded from gene 49.	87	94
448	X14329	Homo sapiens	carboxypeptidase N precursor (AA -20 to 438)	2463	99
449	M36803	Homo sapiens	hemopexin	2603	100
450	AF116238	Homo sapiens	pseudouridine synthase 1	1927	99
451	AB031051	Homo sapiens	organic anion transporter OATP-E	444	42
452	X16841	Homo sapiens	precursor protein (-19 to 742)	3958	100
453	AK022830	Homo sapiens	unnamed protein product	373	100

TABLE 2

SEQ ID NO: OF NUCLEOTIDE	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH- WATERMAN SCORE	% IDENTITY
454	Y94890	Homo sapiens	Human protein clone HP02798.	637	90
455	AL356014	Arabidopsis thaliana	putative protein	210	38
456	X60221	Homo sapiens	H+-ATP synthase subunit b	1297	99
457	G02532	Homo sapiens	Human secreted protein, SEQ ID NO: 6613.	168	69
458	AJ245375	Homo sapiens	PP35 act	1895	99
459	G00397	Homo sapiens	Human secreted protein, SEQ ID NO: 4478.	57	52
460	AE003708	Drosophila melanogaster	CG6194 gene product	234	65
461	W48352	Homo sapiens	Human breast cancer related protein BCFLT1.	80	60
462	U53420	Rattus norvegicus	sodium-calcium exchanger form 3	397	76
463	Y13402	Homo sapiens	Amino acid sequence of protein PRO310.	1075	63
464	Y27607	Homo sapiens	Human secreted protein encoded by gene No. 41.	610	100
465	L08666	Homo sapiens	porin	122	51
466	Y87084	Homo sapiens	Human secreted protein sequence SEQ ID NO:123.	232	78
467	X16841	Homo sapiens	precursor protein (-19 to 742)	3958	100
468	Y48507	Homo sapiens	Human breast tumour-associated protein 52.	295	91
469	X07973	Ovis aries	MT-1b protein	84	45
470	W48927	Homo sapiens	Schwanommin-binding protein C-terminal fragment.	78	60
471	AJ224171	Homo sapiens	lipophilin A	454	100
472	G01984	Homo sapiens	Human secreted protein, SEQ ID NO: 6065.	211	64
473	G03793	Homo sapiens	Human secreted protein, SEQ ID NO: 7874.	200	74
474	Y17829	Homo sapiens	Human PRO354 protein sequence.	1006	100
475	Y66706	Homo sapiens	Membrane-bound protein PRO1129.	2153	99
476	G03800	Homo sapiens	Human secreted protein, SEQ ID NO: 7881.	99	78
477	AF216389	Homo sapiens	semaphorin Rs	296	85

TABLE 2

SEQ ID NO: OF NUCLEOTIDE	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH- WATERMAN SCORE	% IDENTITY
478	X93036	Homo sapiens	MAT8 protein	469	100
479	X53795	Homo sapiens	inducible membrane protein	1412	100
480	AF056195	Homo sapiens	neuroblastoma-amplified protein	4504	98
481	AF116715	Homo sapiens	PRO2829	96	46
482	Z24680	Homo sapiens	garp	167	43
483	Y76198	Homo sapiens	Human secreted protein encoded by gene 75.	82	80
484	AF010144	Homo sapiens	neuronal thread protein AD7c-NTP	324	59
485	Y91592	Homo sapiens	Human secreted protein sequence encoded by gene 6	738	100
486	Y94890	Homo sapiens	Human protein clone HP02798.	605	81
487	U89436	Homo sapiens	tyrosyl-tRNA synthetase	2719	100
488	W88579	Homo sapiens	Secreted protein encoded by gene 46 clone HCFMV39.	479	95
489	G02360	Homo sapiens	Human secreted protein, SEQ ID NO: 6441.	102	70
490	U70976	Homo sapiens	arrestin	1071	61
491	U80746	Homo sapiens	CAGH4	277	81
492	U26361	Helicobacter pylori	Hpn	80	83
493	Y19730	Homo sapiens	SEQ ID NO 448 from WO9922243.	135	53
494	Y27868	Homo sapiens	Human secreted protein encoded by gene No. 107.	185	50
495	AF090901	Homo sapiens	PRO0195	90	46
496	AF061529	Mus musculus	rjs	270	76
497	L34049	Rattus norvegicus	megalin	322	41
498	J04204	Bos taurus	32 kd accessory protein	1743	100
499	Y71118	Homo sapiens	Human Hydrolase protein-16 (HYDRL-16).	2205	97
500	X13916	Homo sapiens	LDL-receptor related precursor (AA -19 to 4525)	715	92
501	Y00877	Homo sapiens	Human LAPH-2 protein sequence.	138	40
502	Y99368	Homo sapiens	Human PRO1326 (UNQ686) amino acid sequence SEQ ID NO:100.	156	48

TABLE 2

SEQ ID NO: OF NUCLEOTIDE	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH- WATERMAN SCORE	% IDENTITY
503	Y48308	Homo sapiens	Human prostate cancer-associated protein 5.	901	100
504	U67060	Cricetulus griseus	SREBP cleavage activating protein	6196	92
505	W75857	Homo sapiens	Human secretory protein of clone CO1020-1.	1761	99
506	X55764	Homo sapiens	11beta-hydrolase precursor	2604	99
507	Y41685	Homo sapiens	Human PRO213 protein sequence.	1344	94
508	X95240	Homo sapiens	cysteine-rich secretory protein-3	1368	100
509	AF065482	Homo sapiens	sorting nexin 2	517	77
510	AF135025	Homo sapiens	kallikrein-like protein 5-related protein 1	1301	100
511	AF220492	Homo sapiens	krueppel-like zinc finger protein HZF2	4100	99
512	X58397	Homo sapiens	variable region V251 from V(H)5 gene	670	100
513	W95348	Homo sapiens	Human foetal kidney secreted protein em397_2.	406	90
514	AJ000479	Homo sapiens	putative G-Protein coupled receptor, EDG6	1966	100
515	L05514	Homo sapiens	histatin 3	280	100
516	X95240	Homo sapiens	cysteine-rich secretory protein-3	1368	100
517	D00654	Homo sapiens	enteric smooth muscle gamma-actin	1972	100
518	AJ005453	Mytilus edulis	metallothionein 10 II	94	35
519	W37864	Homo sapiens	Human protein comprising secretory signal amino acid sequence 1.	362	98
520	X76091	Homo sapiens	DNA binding protein RFX2	3743	99
521	G03800	Homo sapiens	Human secreted protein, SEQ ID NO: 7881.	113	39
522	AJ289243	Mus musculus	calpain 12	147	53
523	D30037	Homo sapiens	phosphatidylinositol transfer protein	1464	100
524	AJ012370	Homo sapiens	NAALADase II protein	3872	99
525	G03909	Homo sapiens	Human secreted protein, SEQ ID NO:	80	41

TABLE 2

SEQ ID NO: OF NUCLEOTIDE	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH- WATERMAN SCORE	% IDENTITY
			7990.		
526	U67060	Cricetulus griseus	SREBP cleavage activating protein	6196	92
527	W48351	Homo sapiens	Human breast cancer related protein BCRB2.	85	61
528	AF093408	Homo sapiens	protein kinase A binding protein AKAP110	461	78
529	Y92182	Homo sapiens	Human partial TANGO 195 from clone T195Athpb93f1.	1682	100
530	M28200	Homo sapiens	MHC class II lymphocyte antigen beta chain	432	72
531	X58397	Homo sapiens	variable region V251 from V(H)5 gene	491	74
532	D88577	Mus musculus	Kupffer cell receptor	904	46
533	M84379	Homo sapiens	lymphocyte antigen	1922	97
534	AF279265	Homo sapiens	putative anion transporter 1	212	91
535	AF132035	Homo sapiens	core 2 beta-1,6-N- acetylglucosaminylt ransferase 3	852	92
536	G02958	Homo sapiens	Human secreted protein, SEQ ID NO: 7039.	512	98
537	Y07938	Homo sapiens	Human secreted protein fragment encoded from gene 87.	302	100
538	Y36203	Homo sapiens	Human secreted protein #75.	175	51
539	U16738	Homo sapiens	CAG-isl 7	472	75
540	AL161531	Arabidopsis thaliana	putative proline- rich protein	118	57
541	K00558	Homo sapiens	alpha-tubulin	2393	100
542	U20286	Rattus norvegicus	lamina associated polypeptide 1C	641	55
543	Y27907	Homo sapiens	Human secreted protein encoded by gene No. 119.	128	61
544	AF109674	Rattus norvegicus	late gestation lung protein 1	954	87
545	L35278	Homo sapiens	bone morphogenetic protein	92	40
546	G00541	Homo sapiens	Human secreted protein, SEQ ID NO: 4622.	94	68
547	AF190664	Mus musculus	LMBR2	246	78
548	Y12793	Homo sapiens	Human 5' EST	113	50

TABLE 2

SEQ ID NO: OF NUCLEOTIDE	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH- WATERMAN SCORE	% IDENTITY
			secreted protein SEQ ID NO:383.		
549	AF133816	Homo sapiens	insulin-like peptide INSL5	714	100
550	X70910	Homo sapiens	ttranectin	1069	100
551	M11902	Mus musculus	proline-rich salivary protein	135	39
552	G03477	Homo sapiens	Human secreted protein, SEQ ID NO: 7558.	89	58
553	U63542	Homo sapiens	FAP protein	156	77
554	Y60497	Homo sapiens	Human normal bladder tissue EST encoded protein 169.	89	50
555	Y87303	Homo sapiens	Human signal peptide containing protein HSPP-80 SEQ ID NO:80.	275	100
556	Y17526	Homo sapiens	Human secreted protein clone AM349 2 protein.	1220	100
557	G04064	Homo sapiens	Human secreted protein, SEQ ID NO: 8145.	83	35
558	U51919	Rattus norvegicus	preprocortistatin	84	36
559	AF090901	Homo sapiens	PRO0195	92	66
560	J04031	Homo sapiens	MDMCSF (EC 1.5.1.5; EC 3.5.4.9; EC 6.3.4.3)	226	52
561	AL117237	Homo sapiens	hypothetical protein	4088	94
562	Y50931	Homo sapiens	Human fetal brain cDNA clone vc25_1 derived protein.	485	100
563	Y21631	Homo sapiens	Ligand binding domain of nuclear receptor hTRbeta.	1738	99
564	X90857	Homo sapiens	-14	177	69
565	W35904	Homo sapiens	Human haematopoietic- specific protein (HSP).	862	87
566	W99070	Homo sapiens	Human PIGR-1.	244	90
567	X61653	Homo sapiens	TCR V-beta 13.5	600	100
568	AF166350	Homo sapiens	ST7 protein	4711	99
569	Y07938	Homo sapiens	Human secreted protein fragment encoded from gene 87.	302	100
570	X85019	Homo sapiens	UDP- GalNAc:polypeptide	3069	100

TABLE 2

SEQ ID NO: OF NUCLEOTIDE	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH- WATERMAN SCORE	% IDENTITY
			N-acetylgalactosaminyl transferase		
571	U89942	Homo sapiens	lysyl oxidase-related protein	2427	89
572	X04391	Homo sapiens	put. precursor polypeptide	2671	99
573	W36903	Homo sapiens	Human epididymis-specific receptor protein.	5352	100
574	U22816	Homo sapiens	LAR-interacting protein 1b	2042	57
575	Y58618	Homo sapiens	Protein regulating gene expression PRGE-11.	729	57
576	AJ278348	Homo sapiens	pregnancy-associated plasma protein-E	743	100
577	AK024512	Homo sapiens	unnamed protein product	471	100
578	AL031685	Homo sapiens	dJ963K23.4 (KIAA0939 (novel Sodium/hydrogen exchanger family member))	2010	100
579	AF183183	Mus musculus	cochlear otoferlin	116	91
580	W74722	Homo sapiens	Human secreted protein er80_1.	2422	100
581	G03356	Homo sapiens	Human secreted protein, SEQ ID NO: 7437.	114	44
582	Y82777	Homo sapiens	Human chordin related protein (Clone dw665_4).	610	98
583	J04988	Homo sapiens	90 kD heat shock protein	3702	100
584	K02576	Homo sapiens	salivary proline-rich protein 1	97	34
585	G03786	Homo sapiens	Human secreted protein, SEQ ID NO: 7867.	159	72
586	AK024490	Homo sapiens	FLJ00092 protein	204	57
587	U22231	Felis catus	ribosomal protein S3a	327	57
588	X55681	Lycopersicon esculentum	extensin (class I)	96	38
589	U68137	Rana ridibunda	prepro-somatostatin 14	81	33
590	Y19655	Homo sapiens	SEQ ID NO 373 from WO9922243.	814	84
591	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	222	56

TABLE 2

SEQ ID NO: OF NUCLEOTIDE	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH- WATERMAN SCORE	% IDENTITY
592	AF067801	Homo sapiens	HDCGC21P	116	38
593	X67339	Neurospora crassa	ccg-2	82	37
594	G03280	Homo sapiens	Human secreted protein, SEQ ID NO: 7361.	169	100
595	Y02693	Homo sapiens	Human secreted protein encoded by gene 44 clone HTDAD22.	130	70
596	AE003683	Drosophila melanogaster	CG9492 gene product	247	56
597	Z22968	Homo sapiens	M130 antigen	6205	100
598	AK021847	Homo sapiens	unnamed protein product	178	94
599	AP000060	Aeropyrum pernix	134aa long hypothetical protein	80	39
600	AK001363	Homo sapiens	unnamed protein product	558	92
601	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	147	49
602	G02538	Homo sapiens	Human secreted protein, SEQ ID NO: 6619.	149	65
603	X98330	Homo sapiens	ryanodine receptor 2	25918	99
604	AJ243460	Leishmania major	proteophosphoglycan	172	35
605	Y81807	Homo sapiens	Human mahogany protein sequence #2.	2499	63
606	AF041069	Equus caballus	fibronectin	109	56
607	Y54591	Homo sapiens	Amino acid sequence of a human transferase designated HUTRAN- 1.	153	77
608	G03172	Homo sapiens	Human secreted protein, SEQ ID NO: 7253.	82	66
609	Y31730	Homo sapiens	Human fused protein kinase-deletion mutant fused C- term.	561	99
610	Y30163	Homo sapiens	Human dorsal root receptor 5 hDRR5.	112	49
611	G03714	Homo sapiens	Human secreted protein, SEQ ID NO: 7795.	171	70
612	U58514	Homo sapiens	chitinase precursor	402	75

TABLE 2

SEQ ID NO: OF NUCLEOTIDE	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH- WATERMAN SCORE	% IDENTITY
613	AL122105	Homo sapiens	hypothetical protein	399	73
614	AF059198	Homo sapiens	protein kinase/endoribonuclease	5093	99
615	X17531	Strongylocentrotus purpuratus	epidermal growth factor	234	54
616	AF112982	Homo sapiens	group IID secretory phospholipase A2	852	100
617	AJ006119	Homo sapiens	anti-IFN-G scFv	675	97
618	W54097	Homo sapiens	Homo sapiens B223 sequence.	339	98
619	AF090930	Homo sapiens	PRO0478	141	79
620	W61624	Homo sapiens	Clone HHFEK40 of TM4SF superfamily.	564	98
621	AF119851	Homo sapiens	PRO1722	115	52
622	G03172	Homo sapiens	Human secreted protein, SEQ ID NO: 7253.	173	48
623	Y41379	Homo sapiens	Human secreted protein encoded by gene 72 clone HE6GA29.	261	100
624	U86339	Drosophila grimshawi	expanded	142	36
625	D86853	Catharanthus roseus	extensin	142	39
626	S58722	Homo sapiens	X-linked retinopathy protein {C-terminal, clone XEH.8c}	116	49
627	G02532	Homo sapiens	Human secreted protein, SEQ ID NO: 6613.	108	50
628	G03790	Homo sapiens	Human secreted protein, SEQ ID NO: 7871.	129	61
629	Y27665	Homo sapiens	Human secreted protein encoded by gene No. 99.	345	100
630	G02837	Homo sapiens	Human secreted protein, SEQ ID NO: 6918.	78	75
631	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	172	65
632	X14329	Homo sapiens	carboxypeptidase N precursor (AA -20 to 438)	2463	99
633	Y87235	Homo sapiens	Human signal peptide containing protein HSPP-12 SEQ	867	100

TABLE 2

SEQ ID NO: OF NUCLEOTIDE	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH- WATERMAN SCORE	% IDENTITY
			ID NO:12.		
634	W88627	Homo sapiens	Secreted protein encoded by gene 94 clone HPMBQ32.	106	73
635	W74845	Homo sapiens	Human secreted protein encoded by gene 117 clone HBMUW78.	395	71
636	M16941	Homo sapiens	DR7 beta-chain glycoprotein	1412	100
637	W95634	Homo sapiens	Homo sapiens secreted protein.	1391	100
638	Y78801	Homo sapiens	Hydrophobic domain containing protein clone HP00631 amino acid sequence.	1277	100
639	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	191	76
640	W64535	Homo sapiens	Human leukocyte cell clone HP00804 protein.	2014	99
641	Y94621	Homo sapiens	Epidermal growth factor-like variant in skin-2 amino acid sequence.	529	91
642	G03646	Homo sapiens	Human secreted protein, SEQ ID NO: 7727.	81	42
643	Y87328	Homo sapiens	Human signal peptide containing protein HSPP-105 SEQ ID NO:105.	681	100
644	Y21386	Homo sapiens	Human HUPF-I mutant protein fragment 34.	78	31
645	G03790	Homo sapiens	Human secreted protein, SEQ ID NO: 7871.	140	55
646	Y35894	Homo sapiens	Extended human secreted protein sequence, SEQ ID NO. 143.	349	100
647	G00517	Homo sapiens	Human secreted protein, SEQ ID NO: 4598.	109	37
648	Y25716	Homo sapiens	Human secreted protein encoded from gene 6.	339	39
649	G01246	Homo sapiens	Human secreted protein, SEQ ID NO: 5327.	152	80
650	R95913	Homo sapiens	Neural thread	233	50

TABLE 2

SEQ ID NO: OF NUCLEOTIDE	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH- WATERMAN SCORE	% IDENTITY
			protein.		
651	Y91469	Homo sapiens	Human secreted protein sequence encoded by gene 19 SEQ ID NO:142.	98	48
652	G03136	Homo sapiens	Human secreted protein, SEQ ID NO: 7217.	94	43
653	U14635	Caenorhabditis elegans	weak similarity to NADH dehydrogenase	186	30
654	Y14482	Homo sapiens	Fragment of human secreted protein encoded by gene 17.	163	54
655	U14635	Caenorhabditis elegans	weak similarity to NADH dehydrogenase	186	30
656	AB024565	Mus musculus	heparan sulfate 6-sulfotransferase 2	1128	79
657	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	243	70
658	Y14471	Homo sapiens	Fragment of human secreted protein encoded by gene 4.	95	65
659	AF135381	Homo sapiens	chemokine-like factor 3	89	59
660	U40407	synthetic construct	T cell receptor alpha chain	586	100
661	AF039712	Caenorhabditis elegans	contains similarity to CDP-alcohol phosphotransferases	289	43
662	G03790	Homo sapiens	Human secreted protein, SEQ ID NO: 7871.	113	55
663	AF084467	Homo sapiens	heparanase	170	32
664	AF279890	Homo sapiens	2P domain potassium channel TREK2	1189	94
665	W63693	Homo sapiens	Human secreted protein 13.	243	84
666	AE003908	Xylella fastidiosa	hypothetical protein	120	28
667	B08948	Homo sapiens	Human secreted protein sequence encoded by gene 21 SEQ ID NO:105.	985	89
668	AF023158	Homo sapiens	tyrosine phosphatase	346	64
669	AF169257	Homo sapiens	sodium/calcium exchanger NCKX3	189	57
670	AF132969	Homo sapiens	CGI-35 protein	364	69
671	AF269286	Homo sapiens	HC6	112	50
672	X98494	Homo sapiens	M phase phosphoprotein 10	529	68
673	G03787	Homo sapiens	Human secreted	83	44

TABLE 2

SEQ ID NO: OF NUCLEOTIDE	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH- WATERMAN SCORE	% IDENTITY
			protein, SEQ ID NO: 7868.		
674	AF119855	Homo sapiens	PRO1847	123	46
675	AJ242540	Volvox carteri f. nagariensis	hydroxyproline-rich glycoprotein DZ-HRGP	242	42
676	Y91666	Homo sapiens	Human secreted protein sequence encoded by gene 72 SEQ ID NO:339.	529	96
677	Y57936	Homo sapiens	Human transmembrane protein HTMPN-60.	669	100
678	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	156	72
679	W18878	Homo sapiens	Human protein kinase C inhibitor, IPKC-1.	98	68
680	Z12168	Canis familiaris	stimulatory GTP binding protein	980	88
681	G00517	Homo sapiens	Human secreted protein, SEQ ID NO: 4598.	160	48
682	W19932	Homo sapiens	Alzheimer's disease protein encoded by DNA from plasmid pGCS55.	362	100
683	Y30709	Homo sapiens	Amino acid sequence of a human secreted protein.	99	56
684	AF269286	Homo sapiens	HC6	137	72
685	M14362	Homo sapiens	T-cell surface antigen CD2 precursor	275	64
686	G02493	Homo sapiens	Human secreted protein, SEQ ID NO: 6574.	173	61
687	AF248635	Mus musculus	lymphocyte antigen 108 isoform 1	303	50
688	D86983	Homo sapiens	similar to D.melanogaster peroxidasin (U11052)	288	55
689	Y59711	Homo sapiens	Secreted protein 58-20-4-G7-FL1.	895	91
690	W48848	Homo sapiens	Human receptor tyrosine kinase LMR3_h N-terminal polypeptide.	1056	89
691	W22652	Homo sapiens	64-863 antibody HSV863 light chain variable region.	459	77
692	AF098066	Homo sapiens	squamous cell carcinoma antigen	1001	98

TABLE 2

SEQ ID NO: OF NUCLEOTIDE	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH- WATERMAN SCORE	% IDENTITY
			recognized by T cell		
693	D83039	Homo sapiens	eti-1	426	98
694	Y79511	Homo sapiens	Human carbohydrate-associated protein CRBAP-7.	1245	99
695	U12623	Rattus norvegicus	cyclic nucleotide gated cation channel	857	83
696	AF229067	Homo sapiens	PADI-H protein	174	61
697	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	196	75
698	U10921	Macaca mulatta	T-cell receptor alpha chain	578	82
699	U31913	Homo sapiens	HBV-X associated protein	167	100
700	X99043	Mus musculus	brain-derived immunoglobulin superfamily molecule	348	82
701	X59770	Homo sapiens	type II interleukin-1 receptor	2130	100
702	AC018758	Homo sapiens	GPI-anchored metastasis-associated protein homolog	207	31
703	Y28816	Homo sapiens	pm4_13 secreted protein.	280	100
704	Y52386	Homo sapiens	Human transmembrane protein HP02000.	1077	100
705	U12392	Haematobia irritans	putative ATPase	481	55
706	U11265	Homo sapiens	HLA-B35	351	92
707	X64594	Homo sapiens	50 kDa erythrocyte plasma membrane glycoprotein	301	88
708	AB046048	Macaca fascicularis	unnamed protein product	260	67
709	G03807	Homo sapiens	Human secreted protein, SEQ ID NO: 7888.	119	60
710	G03315	Homo sapiens	Human secreted protein, SEQ ID NO: 7396.	314	100
711	Y50945	Homo sapiens	Human adult thymus cDNA clone vh1_1 derived protein #1.	742	100
712	G00564	Homo sapiens	Human secreted protein, SEQ ID NO: 4645.	271	98
713	G00125	Homo sapiens	Human secreted	373	80

TABLE 2

SEQ ID NO: OF NUCLEOTIDE	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH- WATERMAN SCORE	% IDENTITY
			protein, SEQ ID NO: 4206.		
714	Y13352	Homo sapiens	Amino acid sequence of protein PRO228.	872	98
715	G02753	Homo sapiens	Human secreted protein, SEQ ID NO: 6834.	222	68
716	Y19588	Homo sapiens	Amino acid sequence of a human secreted protein.	329	100
717	AB030235	Canis familiaris	D4 dopamine receptor	79	35
718	W74577	Homo sapiens	Human membrane protein BA2303.	748	100
719	Y02693	Homo sapiens	Human secreted protein encoded by gene 44 clone HTDAD22.	235	61
720	X97868	Homo sapiens	arylsulphatase	167	84
721	Y13215	Homo sapiens	Human secreted protein encoded by 5' EST SEQ ID NO: 229.	234	97
722	Y20298	Homo sapiens	Human apolipoprotein E mutant protein fragment 11.	152	39
723	Y86231	Homo sapiens	Human secreted protein HLTHR66, SEQ ID NO:146.	207	51
724	W75083	Homo sapiens	Human secreted protein encoded by gene 27 clone HSPAF93.	685	100
725	W88627	Homo sapiens	Secreted protein encoded by gene 94 clone HPMBQ32.	301	73
726	Y27868	Homo sapiens	Human secreted protein encoded by gene No. 107.	229	58
727	AK025470	Homo sapiens	unnamed protein product	130	64
728	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	159	46
729	Y25776	Homo sapiens	Human secreted protein encoded from gene 66.	334	43
730	AF116661	Homo sapiens	PRO1438	153	56
731	W48351	Homo sapiens	Human breast cancer related protein BCRB2.	106	72
732	U77589	Homo sapiens	MHC class II HLA-	133	69

TABLE 2

SEQ ID NO: OF NUCLEOTIDE	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH- WATERMAN SCORE	% IDENTITY
			DQ-alpha chain		
733	G00357	Homo sapiens	Human secreted protein, SEQ ID NO: 4438.	223	67
734	R28542	Homo sapiens	Human complement type 1 receptor SCR9.	152	96
735	Y27868	Homo sapiens	Human secreted protein encoded by gene No. 107.	150	65
736	AB036706	Homo sapiens	intelectin	368	76
737	Y74042	Homo sapiens	Human prostate tumor EST fragment derived protein #229.	206	65
738	Y36156	Homo sapiens	Human secreted protein #28.	153	77
739	W74802	Homo sapiens	Human secreted protein encoded by gene 73 clone HSQEL25.	1751	79
740	W85614	Homo sapiens	Secreted protein clone fr473_2.	224	91
741	Y13377	Homo sapiens	Amino acid sequence of protein PRO257.	394	98
742	Z69384	Caenorhabditis elegans	Similarity to Salmonella regulatory protein UHPC (SW:UHPC_SALTY)	515	45
743	W47589	Homo sapiens	T-cell receptor beta-chain.	681	92
744	G03786	Homo sapiens	Human secreted protein, SEQ ID NO: 7867.	243	71
745	Y50690	Homo sapiens	Human Hum4 VL ClaI-HindIII segment encoded protein.	540	81
746	U03414	Rattus norvegicus	neuronal olfactomedin-related ER localized protein	363	67
747	G00352	Homo sapiens	Human secreted protein, SEQ ID NO: 4433.	84	51
748	Y02671	Homo sapiens	Human secreted protein encoded by gene 22 clone HMSJW18.	145	60
749	AF026919	Homo sapiens	amyloid lambda light chain variable region	557	83
750	X76732	Homo sapiens	NEFA protein	297	100

TABLE 2

SEQ ID NO: OF NUCLEOTIDE	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH- WATERMAN SCORE	% IDENTITY
751	R92754	Homo sapiens	Human growth differentiation factor-12.	628	100
752	Y91462	Homo sapiens	Human secreted protein sequence encoded by gene 12 SEQ ID NO:135.	597	100
753	Y66700	Homo sapiens	Membrane-bound protein PRO1137.	754	99
754	G01648	Homo sapiens	Human secreted protein, SEQ ID NO: 5729.	281	100
755	AB040434	Homo sapiens	hTROY	752	100
756	Y28680	Homo sapiens	Human nm214_3 secreted protein.	178	44
757	W75100	Homo sapiens	Human secreted protein encoded by gene 44 clone HE8CJ26.	203	66
758	AF090930	Homo sapiens	PRO0478	87	45
759	D84336	Rattus norvegicus	ZOG	484	48
760	W88627	Homo sapiens	Secreted protein encoded by gene 94 clone HPMBQ32.	150	81
761	Y48616	Homo sapiens	Human breast tumour-associated protein 77.	569	70
762	Y87320	Homo sapiens	Human signal peptide containing protein HSPP-97 SEQ ID NO:97.	918	100
763	G03655	Homo sapiens	Human secreted protein, SEQ ID NO: 7736.	248	89
764	AF031174	Homo sapiens	Ig-like membrane protein	428	45
765	U08255	Rattus norvegicus	glutamate receptor delta-1 subunit	802	99
766	Y99369	Homo sapiens	Human PRO1249 (UNQ632) amino acid sequence SEQ ID NO:102.	4578	99
767	AK001586	Homo sapiens	unnamed protein product	973	98
768	AC007063	Arabidopsis thaliana	putative ABC transporter	126	31
769	AF303378	Homo sapiens	sialic acid-specific acetylesterase II	713	100
770	G00517	Homo sapiens	Human secreted protein, SEQ ID NO: 4598.	90	37

TABLE 2

SEQ ID NO: OF NUCLEOTIDE	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH- WATERMAN SCORE	% IDENTITY
771	Y59733	Homo sapiens	Human normal ovarian tissue derived protein 10.	1253	99
772	AF132856	Homo sapiens	suppressor of G2 allele of skp1 homolog	163	86
773	AB029482	Mus musculus	JNK-binding protein JNKBP1	1082	97
774	G02108	Homo sapiens	Human secreted protein, SEQ ID NO: 6189.	134	62
775	AB047818	Homo sapiens	Scggy	1239	100
776	Y66689	Homo sapiens	Membrane-bound protein PRO1136.	804	99
777	Y71107	Homo sapiens	Human Hydrolase protein-5 (HYDRL-5).	733	99
778	AC005626	Homo sapiens	R29124_1	182	38
779	W88707	Homo sapiens	Secreted protein encoded by gene 174 clone HE9FB42.	126	56
780	G03657	Homo sapiens	Human secreted protein, SEQ ID NO: 7738.	455	96
781	AJ001616	Mus musculus	myeloid associated differentiation protein	201	36
782	Y64942	Homo sapiens	Human 5' EST related polypeptide SEQ ID NO:1103.	86	65
783	AL356276	Homo sapiens	bA367J7.2.1 (novel Immunoglobulin domains containing protein (isoform 1))	845	91
784	Y00876	Homo sapiens	Human LAPH-1 protein sequence.	291	43
785	G00270	Homo sapiens	Human secreted protein, SEQ ID NO: 4351.	603	100
786	AF154121	Homo sapiens	sodium-dependent high-affinity dicarboxylate transporter	864	100
787	Y29804	Homo sapiens	Human GABA B receptor subunit HG20 peptide #6.	83	42
788	AL080239	Homo sapiens	bG256O22.1 (similar to IGFALS (insulin-like growth factor binding protein, acid labile subunit))	599	100

TABLE 2

SEQ ID NO: OF NUCLEOTIDE	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH- WATERMAN SCORE	% IDENTITY
789	AL031856	<i>Schizosaccharomyces pombe</i>	PUTATIVE GOLGI URIDINE DIPHOSPHATE-N-ACETYLGLUCOSAMINE TRANSPORTER	192	40
790	G03448	<i>Homo sapiens</i>	Human secreted protein, SEQ ID NO: 7529.	141	43
791	U81291	<i>Xenopus laevis</i>	oviductin	310	38
792	Y41332	<i>Homo sapiens</i>	Human secreted protein encoded by gene 25 clone HPIBO48.	295	50
793	L20315	<i>Mus musculus</i>	MPS1 protein	702	77
794	G01314	<i>Homo sapiens</i>	Human secreted protein, SEQ ID NO: 5395.	91	36
795	AF003136	<i>Caenorhabditis elegans</i>	similar to 1-acyl-glycerol-3-phosphate acyltransferases	122	38
796	G00637	<i>Homo sapiens</i>	Human secreted protein, SEQ ID NO: 4718.	160	67
797	Y36144	<i>Homo sapiens</i>	Human secreted protein #16.	622	100
798	U09453	<i>Cricetulus griseus</i>	UDP-N-acetylglucosamine: dolichyl phosphate N-acetylglucosamine 1-phosphate transferase	178	66
799	Y76144	<i>Homo sapiens</i>	Human secreted protein encoded by gene 21.	633	100
800	Y73456	<i>Homo sapiens</i>	Human secreted protein clone yd145_1 protein sequence SEQ ID NO:134.	413	89
801	Y86540	<i>Homo sapiens</i>	Human gene 77-encoded protein fragment, SEQ ID NO:457.	443	96
802	U49973	<i>Homo sapiens</i>	ORF1; MER37; putative transposase similar to pogo element	311	53
803	M63573	<i>Homo sapiens</i>	secreted cyclophilin-like protein	700	88
804	AF091622	<i>Homo sapiens</i>	PHD finger protein	177	100

TABLE 2

SEQ ID NO: OF NUCLEOTIDE	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH- WATERMAN SCORE	% IDENTITY
			3		
805	W37869	Homo sapiens	Human protein comprising secretory signal amino acid sequence 6.	381	100
806	G03556	Homo sapiens	Human secreted protein, SEQ ID NO: 7637.	221	72
807	AF178941	Homo sapiens	ATP-binding cassette sub-family A member 2	583	87
808	Y91385	Homo sapiens	Human secreted protein sequence encoded by gene 40 SEQ ID NO:106.	786	100
809	Y00826	Rattus norvegicus	gp210 (AA 1-1886)	169	83
810	G03143	Homo sapiens	Human secreted protein, SEQ ID NO: 7224.	328	100
811	W00870	Homo sapiens	Polycystic kidney disease 1 (PKD1) polypeptide.	22446	99
812	Y73434	Homo sapiens	Human secreted protein clone yd51_1 protein sequence SEQ ID NO: 90.	417	90
813	AB031996	Ralstonia sp. KN1	ferredoxin-like protein	94	44
814	AF201734	Mus musculus	testis specific serine kinase-3	800	87
815	Y01181	Homo sapiens	Polypeptide fragment encoded by gene 12.	68	55
816	Y76166	Homo sapiens	Human secreted protein encoded by gene 43.	724	94
817	AL109827	Homo sapiens	dJ309K20.2 (acrosomal protein ACR55 (similar to rat sperm antigen 4 (SPAG4)))	639	84
818	M62829	Homo sapiens	ETR103	137	53
819	Y38422	Homo sapiens	Human secreted protein.	526	100
820	AF119815	Homo sapiens	G-protein-coupled receptor	561	79
821	Y87101	Homo sapiens	Human secreted protein sequence SEQ ID NO:140.	628	100
822	M91463	Homo sapiens	glucose transporter	213	79

TABLE 2

SEQ ID NO: OF NUCLEOTIDE	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH- WATERMAN SCORE	% IDENTITY
823	L34938	Rattus norvegicus	ionotropic glutamate receptor	618	90
824	W17846	Homo sapiens	Cytosolic phospholipase A2/B (clone 19b product).	209	64
825	Y66722	Homo sapiens	Membrane-bound protein PRO1104.	221	67
826	G02493	Homo sapiens	Human secreted protein, SEQ ID NO: 6574.	138	72
827	Y91423	Homo sapiens	Human secreted protein sequence encoded by gene 11 SEQ ID NO:144.	671	54
828	U78090	Rattus norvegicus	potassium channel regulator 1	502	80
829	U08813	Oryctolagus cuniculus	597 aa protein related to Na/glucose cotransporters	906	84
830	AJ272063	Homo sapiens	vanilloid receptor 1	630	90
831	U36898	Rattus norvegicus	pheromone receptor VN6	135	52
832	Z46973	Homo sapiens	phosphatidylinosito l 3-kinase	396	80
833	Y95433	Homo sapiens	Human calcium channel SOC-2/CRAC- 1 C-terminal polypeptide.	747	99
834	AF132856	Homo sapiens	suppressor of G2 allele of skp1 homolog	163	86
835	AC006042	Homo sapiens	supported by human ESTs AI681256.1(NID:g489 1438), N32168.1(NID: g1152567), and genscan	195	87
836	B01247	Homo sapiens	Human HE6 receptor.	371	45
837	G03788	Homo sapiens	Human secreted protein, SEQ ID NO: 7869.	196	59
838	U70136	Homo sapiens	megakaryocyte stimulating factor; MSF	6954	98
839	AF017153	Mus musculus	putative RNA helicase and RNA dependent ATPase	178	51
840	Y31830	Homo sapiens	Human adult brain secreted protein nh899_8.	244	56

TABLE 2

SEQ ID NO: OF NUCLEOTIDE	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH- WATERMAN SCORE	% IDENTITY
841	Y27593	Homo sapiens	Human secreted protein encoded by gene No. 27.	437	81
842	G01984	Homo sapiens	Human secreted protein, SEQ ID NO: 6065.	196	74
843	AL008723	Homo sapiens	dJ90G24.4 (SAAT1 (low affinity sodium glucose cotransporter (sodium:solute symporter family)))	183	92
844	AF068065	Cryptosporidium parvum	GP900; mucin-like glycoprotein	263	47
845	Y00815	Homo sapiens	put. LAR preprotein (AA -16 to 1881)	341	100
846	Y06816	Homo sapiens	Human Notch2 (humN2) protein sequence.	1224	99
847	AF104923	Homo sapiens	putative transcription factor	293	95
848	Y09945	Rattus norvegicus	putative integral membrane transport protein	589	53
849	AL157874	Schizosaccharomyces pombe	similar to yeast SCT1 suppressor of a choline transport mutant	146	40
850	R71003	Homo sapiens	Human neuronal calcium channel subunit alpha 1c-1.	141	89
851	X75756	Homo sapiens	protein kinase C mu	318	90
852	AF142676	Drosophila melanogaster	sodium-hydrogen exchanger NHE1	366	48
853	Y45381	Homo sapiens	Human secreted protein fragment encoded from gene 28.	139	73
854	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	121	60
855	U65409	Yarrowia lipolytica	Sla2p	109	25
856	M19419	Mus musculus	proline-rich salivary protein	109	36
857	Y99355	Homo sapiens	Human PRO1295 (UNQ664) amino acid sequence SEQ ID NO:54.	667	98
858	W19919	Homo sapiens	Human Ksr-1 (kinase suppressor of Ras).	211	86
859	Y95436	Homo sapiens	Human calcium	764	84

TABLE 2

SEQ ID NO: OF NUCLEOTIDE	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH- WATERMAN SCORE	% IDENTITY
			channel SOC-3/CRAC-2.		
860	AF070066	Mus musculus	Citron-K kinase	628	97
861	AF286095	Homo sapiens	IL-22 receptor	933	100
862	AF020195	Mus musculus	pancreas sodium bicarbonate cotransporter	475	68
863	G03712	Homo sapiens	Human secreted protein, SEQ ID NO: 7793.	240	100
864	AF195092	Homo sapiens	sialic acid-binding immunoglobulin-like lectin-8	288	87
865	AF208110	Homo sapiens	IL-17 receptor homolog precursor	2688	99
866	L42338	Mus musculus	sodium channel 25	733	98
867	G02360	Homo sapiens	Human secreted protein, SEQ ID NO: 6441.	101	70
868	AF065215	Homo sapiens	cytosolic phospholipase A2 beta	290	42
869	L43631	Homo sapiens	scaffold attachment factor B	106	95
870	G03034	Homo sapiens	Human secreted protein, SEQ ID NO: 7115.	108	54
871	Z21514	Rattus norvegicus	integral membrane glycoprotein	84	47
872	AF097518	Homo sapiens	liver-specific transporter	147	40
873	AF288223	Drosophila melanogaster	Crossveinless 2	136	39
874	U90126	Bos taurus	ABC transporter	245	36
875	AF099988	Mus musculus	Ste-20 related kinase SPAK	103	34
876	Y70400	Homo sapiens	Human cell-signalling protein-2.	220	86
877	Y36300	Homo sapiens	Human secreted protein encoded by gene 77.	1863	99
878	AF151074	Homo sapiens	HSPC240	193	29
879	Y94951	Homo sapiens	Human secreted protein clone dw78_1 protein sequence SEQ ID NO:108.	251	89
880	AF165310	Homo sapiens	ATP cassette binding transporter 1	231	31
881	AF252281	Mus musculus	Kelch-like 1 protein	256	58

TABLE 2

SEQ ID NO: OF NUCLEOTIDE	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH- WATERMAN SCORE	% IDENTITY
882	Y00931	Homo sapiens	Prostate-tumour derived antigen #4.	1039	98
883	Y27576	Homo sapiens	Human secreted protein encoded by gene No. 10.	394	96
884	U00009	Escherichia coli	yeeF	153	30
885	Y57945	Homo sapiens	Human transmembrane protein HTMPN-69.	1543	100
886	Y28678	Homo sapiens	Human cw272_7 secreted protein.	375	60
887	W95349	Homo sapiens	Human foetal brain secreted protein fh170_7.	377	89
888	Y87329	Homo sapiens	Human signal peptide containing protein HSPP-106 SEQ ID NO:106.	285	89
889	AL121845	Homo sapiens	dJ583P15.5.1 (novel protein (isoform 1))	1399	99
890	R75181	Homo sapiens	Partial peptide of human HMW kininogen fragment 1.2.	100	29
891	AF105365	Homo sapiens	K-Cl cotransporter KCC4	680	100
892	Y91644	Homo sapiens	Human secreted protein sequence encoded by gene 43 SEQ ID NO:317.	673	95
893	S52051	Rattus sp.	neurotransmitter transporter	656	99
894	S52051	Rattus sp.	neurotransmitter transporter	617	94
895	R47120	Homo sapiens	Partial human H13 polypeptide.	343	60
896	Z98046	Homo sapiens	dJ1409.2 (Melanoma-Associated Antigen MAGE LIKE)	332	49
897	AJ006203	Oryctolagus cuniculus	capacitative calcium entry channel 2	740	99
898	AF156547	Mus musculus	putative E1-E2 ATPase	769	95
899	AC004076	Homo sapiens	R30217_1	788	98
900	D00099	Homo sapiens	Na,K-ATPase alpha-subunit	753	94
901	R27648	Homo sapiens	Human calcium channel 27980/10.	536	85
902	Y57955	Homo sapiens	Human transmembrane protein HTMPN-79.	606	100
903	AF155913	Mus musculus	putative E1-E2 ATPase	1039	85

TABLE 2

SEQ ID NO: OF NUCLEOTIDE	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH- WATERMAN SCORE	% IDENTITY
904	Y73446	Homo sapiens	Human secreted protein clone yc27_1 protein sequence SEQ ID NO:114.	369	66
905	Y94903	Homo sapiens	Human secreted protein clone pt332_1 protein sequence SEQ ID NO:12.	3777	100
906	AB032470	Homo sapiens	seven transmembrane protein TM7SF3	2124	100
907	G00517	Homo sapiens	Human secreted protein, SEQ ID NO: 4598.	90	50
908	AF010144	Homo sapiens	neuronal thread protein AD7c-NTP	270	65
909	AF263912	Streptomyces noursei	NysA	113	25
910	Y53051	Homo sapiens	Human secreted protein clone dd119_4 protein sequence SEQ ID NO:108.	843	49
911	Y76179	Homo sapiens	Human secreted protein encoded by gene 56.	634	100
912	G00352	Homo sapiens	Human secreted protein, SEQ ID NO: 4433.	229	71
913	U93569	Homo sapiens	p40	110	32
914	G02639	Homo sapiens	Human secreted protein, SEQ ID NO: 6720.	65	46
915	Y94951	Homo sapiens	Human secreted protein clone dw78_1 protein sequence SEQ ID NO:108.	100	38
916	G03263	Homo sapiens	Human secreted protein, SEQ ID NO: 7344.	80	47
917	W74887	Homo sapiens	Human secreted protein encoded by gene 160 clone HCELB21.	273	69
918	Y73464	Homo sapiens	Human secreted protein clone y14_1 protein sequence SEQ ID NO:150.	982	90
919	AF064801	Homo sapiens	multiple membrane spanning receptor TRC8	551	32

TABLE 2

SEQ ID NO: OF NUCLEOTIDE	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH- WATERMAN SCORE	% IDENTITY
920	Y87335	Homo sapiens	Human signal peptide containing protein HSPP-112 SEQ ID NO:112.	622	99
921	AK000496	Homo sapiens	unnamed protein product	342	74
922	Y41360	Homo sapiens	Human secreted protein encoded by gene 53 clone HJPAD75.	367	100
923	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	328	75
924	Y53881	Homo sapiens	A suppressor of cytokine signalling protein designated HSCOP-1.	1489	100
925	AC004144	Homo sapiens	R34001_1	193	60
926	AF119851	Homo sapiens	PRO1722	153	82
927	G02654	Homo sapiens	Human secreted protein, SEQ ID NO: 6735.	82	57
928	Y30819	Homo sapiens	Human secreted protein encoded from gene 9.	264	33
929	G01691	Homo sapiens	Human secreted protein, SEQ ID NO: 5772.	66	43
930	AF187845	Homo sapiens	small protein effector 1 of Cdc42	431	100
931	AL390114	Leishmania major	extremely cysteine/valine rich protein	113	40
932	AL080239	Homo sapiens	bG256O22.1 (similar to IGFLS (insulin-like growth factor binding protein, acid labile subunit))	1451	97
933	W85613	Homo sapiens	Secreted protein clone fm60_1.	234	100
934	AF009243	Homo sapiens	proline-rich Gla protein 2	223	42
935	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	271	66
936	AK000385	Homo sapiens	unnamed protein product	193	64
937	AF010144	Homo sapiens	neuronal thread protein AD7c-NTP	270	65
938	AF119851	Homo sapiens	PRO1722	170	71
939	Y07922	Homo sapiens	Human secreted protein fragment	226	95

TABLE 2

SEQ ID NO: OF NUCLEOTIDE	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH- WATERMAN SCORE	% IDENTITY
			encoded from gene 71.		
940	Y41712	Homo sapiens	Human PRO724 protein sequence.	653	96
941	AF010144	Homo sapiens	neuronal thread protein AD7c-NTP	310	64
942	Y45318	Homo sapiens	Human secreted protein fragment encoded from gene 18.	502	98
943	Y07899	Homo sapiens	Human secreted protein fragment encoded from gene 48.	309	98
944	X92485	Plasmodium vivax	pval	185	51
945	AJ289133	Mus musculus	chondroitin 4-O-sulfotransferase	565	43
946	AF151074	Homo sapiens	HSPC240	1337	99
947	U40829	Saccharomyces cerevisiae	Weak similarity near C-terminus to RNA Polymerase beta subunit (Swiss Prot. accession number P11213) and CCAAT-binding transcription factor (PIR accession number A36368)	361	50
948	Y87285	Homo sapiens	Human signal peptide containing protein HSPP-62 SEQ ID NO:62.	348	82
949	Y86230	Homo sapiens	Human secreted protein HKFBC53, SEQ ID NO:145.	368	80
950	AJ010346	Homo sapiens	RING-H2	333	87
951	Z56281	Homo sapiens	interferon regulatory factor 3	1573	81
952	Y57896	Homo sapiens	Human transmembrane protein HTMPN-20.	421	100
953	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	135	55
954	Y87103	Homo sapiens	Human secreted protein sequence SEQ ID NO:142.	83	50
955	Y87345	Homo sapiens	Human signal peptide containing protein HSPP-122 SEQ ID NO:122.	885	99
956	X81479	Homo sapiens	EMR1	1148	99

TABLE 2

SEQ ID NO: OF NUCLEOTIDE	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH- WATERMAN SCORE	% IDENTITY
957	AF175406	Homo sapiens	transient receptor potential 4	4061	99
958	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	276	73
959	M63274	Plasmodium falciparum	malaria antigen	77	38
960	Y78795	Homo sapiens	Human antizuai-2 (AZ-2) amino acid sequence.	3384	83
961	AL133469	Streptomyces coelicolor A3 (2)	putative secreted proline-rich protein	139	41
962	G03787	Homo sapiens	Human secreted protein, SEQ ID NO: 7868.	232	72
963	W74828	Homo sapiens	Human secreted protein encoded by gene 100 clone HLQAB52.	1016	99
964	W48351	Homo sapiens	Human breast cancer related protein BCRB2.	226	58
965	X63893	Sus scrofa	alpha-stimulatory subunit of GTP-binding protein	319	86
966	AB033019	Homo sapiens	KIAA1193 protein	245	97
967	Y36156	Homo sapiens	Human secreted protein #28.	223	85
968	AF119851	Homo sapiens	PRO1722	188	69
969	Y15224	Homo sapiens	Human receptor protein (HURP) 3 amino acid sequence.	214	42
970	G02754	Homo sapiens	Human secreted protein, SEQ ID NO: 6835.	81	62
971	U22376	Homo sapiens	alternatively spliced product using exon 13A	212	81
972	W74870	Homo sapiens	Human secreted protein encoded by gene 142 clone HTWCB92.	164	81
973	Y30817	Homo sapiens	Human secreted protein encoded from gene 7.	717	98
974	AF079529	Homo sapiens	cAMP-specific phosphodiesterase 8B; PDE8B1; 3',5'-cyclic nucleotide phosphodiesterase	2353	96
975	AF099028	Drosophila	putative	1061	52

TABLE 2

SEQ ID NO: OF NUCLEOTIDE	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH- WATERMAN SCORE	% IDENTITY
		melanogaster	transmembrane protein cmp44E		
976	G03786	Homo sapiens	Human secreted protein, SEQ ID NO: 7867.	179	72
977	Y22495	Homo sapiens	Human secreted protein sequence clone ch4_11.	1629	100
978	W74813	Homo sapiens	Human secreted protein encoded by gene 85 clone HSDFV29.	722	92
979	AK023408	Homo sapiens	unnamed protein product	974	96
980	AF229178	Homo sapiens	leucine rich repeat and death domain containing protein	276	67
981	G03797	Homo sapiens	Human secreted protein, SEQ ID NO: 7878.	198	56
982	W74831	Homo sapiens	Human secreted protein encoded by gene 103 clone HEBDJ82.	153	100
983	G01335	Homo sapiens	Human secreted protein, SEQ ID NO: 5416.	157	96
984	Y73436	Homo sapiens	Human secreted protein clone ye43_1 protein sequence SEQ ID NO: 94.	450	100
985	G00354	Homo sapiens	Human secreted protein, SEQ ID NO: 4435.	96	58
986	Y41712	Homo sapiens	Human PRO724 protein sequence.	639	88
987	Y57896	Homo sapiens	Human transmembrane protein HTMPN-20.	421	100
988	Y66691	Homo sapiens	Membrane-bound protein PRO809.	716	65
989	AF090943	Homo sapiens	PRO0659	926	100
990	G00403	Homo sapiens	Human secreted protein, SEQ ID NO: 4484.	80	46
991	G03411	Homo sapiens	Human secreted protein, SEQ ID NO: 7492.	62	57
992	G00270	Homo sapiens	Human secreted protein, SEQ ID NO: 4351.	143	96
993	AF026246	Homo sapiens	HERV-E integrase	361	80
994	Y36421	Homo sapiens	Fragment of human	83	37

TABLE 2

SEQ ID NO: OF NUCLEOTIDE	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH- WATERMAN SCORE	% IDENTITY
			secreted protein encoded by gene 8.		
995	U22376	Homo sapiens	alternatively spliced product using exon 13A	175	78
996	G03790	Homo sapiens	Human secreted protein, SEQ ID NO: 7871.	87	35
997	G00397	Homo sapiens	Human secreted protein, SEQ ID NO: 4478.	149	61
998	J02642	Homo sapiens	glyceraldehyde 3- phosphate dehydrogenase (EC 1.2.1.12)	429	69
999	AF119851	Homo sapiens	PRO1722	204	50
1000	Y91423	Homo sapiens	Human secreted protein sequence encoded by gene 11 SEQ ID NO:144.	393	53
1001	Y66695	Homo sapiens	Membrane-bound protein PRO1344.	1183	87
1002	AF090931	Homo sapiens	PRO0483	149	68
1003	Y33261	Homo sapiens	Human p99 protein.	314	59
1004	U11494	Mus musculus	protein kinase	360	77
1005	AK021848	Homo sapiens	unnamed protein product	186	69
1006	Y13892	Homo sapiens	PI-3 kinase	233	97
1007	W48351	Homo sapiens	Human breast cancer related protein BCRB2.	144	65
1008	G03793	Homo sapiens	Human secreted protein, SEQ ID NO: 7874.	202	67
1009	U91682	Aedes aegypti	vitelline membrane protein homolog	88	42

TABLE 3

SEQ ID NO: OF NUCLEOTIDE	SEQ ID NO: OF AMINO ACID	SEQ ID NO: IN USSN 09/491,404	START NUCLEOTIDE OF CODING REGION	STOP NUCLEOTIDE OF CODING REGION
1	1010	100	299	535
2	1011	1002	19	267
3	1012	1003	31	423
4	1013	1007	148	840
5	1014	1009	139	318
6	1015	1010	413	748
7	1016	1012	357	154
8	1017	1014	133	285
9	1018	1016	61	441
10	1019	102	269	832
11	1020	1021	148	342
12	1021	1022	45	452
13	1022	1035	222	779
14	1023	1038	222	779
15	1024	1042	735	517
16	1025	1049	120	320
17	1026	1055	195	395
18	1027	1061	13	189
19	1028	1070	972	1109
20	1029	1071	1504	1686
21	1030	1077	425	574
22	1031	108	46	501
23	1032	1088	1949	7240
24	1033	1092	119	571
25	1034	1095	118	564
26	1035	1096	110	373
27	1036	1098	66	353
28	1037	1099	1	417
29	1038	11	764	573
30	1039	1100	157	1014
31	1040	1102	1526	1813
32	1041	1103	1529	1338
33	1042	1104	685	1929
34	1043	1105	887	744
35	1044	1110	880	443
36	1045	1111	696	538
37	1046	1113	52	1272
38	1047	1117	1357	554
39	1048	1118	1478	1654
40	1049	112	482	712
41	1050	1121	3	1424
42	1051	1130	131	271
43	1052	1132	849	151
44	1053	1137	265	705
45	1054	1138	13	381
46	1055	1140	51	416
47	1056	1146	2389	2541
48	1057	1148	1517	738
49	1058	115	179	334
50	1059	1154	68	358

TABLE 3

SEQ ID NO: OF NUCLEOTIDE	SEQ ID NO: OF AMINO ACID	SEQ ID NO: IN USSN 09/491,404	START NUCLEOTIDE OF CODING REGION	STOP NUCLEOTIDE OF CODING REGION
51	1060	1155	34	330
52	1061	1157	242	433
53	1062	1160	410	856
54	1063	1161	154	342
55	1064	1163	202	477
56	1065	1167	72	272
57	1066	117	235	2
58	1067	1170	47	211
59	1068	1176	16	159
60	1069	1177	135	326
61	1070	118	1248	1466
62	1071	1183	431	886
63	1072	1187	191	529
64	1073	1189	1303	1148
65	1074	119	380	613
66	1075	1190	514	1272
67	1076	1192	1529	1338
68	1077	1197	93	533
69	1078	1199	227	391
70	1079	1202	117	407
71	1080	1204	12	413
72	1081	1205	49	603
73	1082	1216	487	1341
74	1083	1217	982	764
75	1084	1228	99	266
76	1085	1230	973	770
77	1086	1233	233	418
78	1087	1234	2959	2078
79	1088	1235	112	1542
80	1089	1239	3019	2822
81	1090	1242	1335	781
82	1091	1248	29	169
83	1092	125	542	405
84	1093	1250	1381	1572
85	1094	1252	480	226
86	1095	1255	19	285
87	1096	1259	165	638
88	1097	126	627	364
89	1098	1260	289	462
90	1099	1262	138	353
91	1100	1264	1159	1299
92	1101	1266	13	402
93	1102	1269	296	805
94	1103	127	212	397
95	1104	1270	126	374
96	1105	1272	2025	2396
97	1106	1273	1367	624
98	1107	1274	1108	746
99	1108	1275	919	1077
100	1109	1279	496	1272

TABLE 3

SEQ ID NO: OF NUCLEOTIDE	SEQ ID NO: OF AMINO ACID	SEQ ID NO: IN USSN 09/491,404	START NUCLEOTIDE OF CODING REGION	STOP NUCLEOTIDE OF CODING REGION
101	1110	1283	265	125
102	1111	1287	107	385
103	1112	1297	333	545
104	1113	13	187	47
105	1114	130	126	290
106	1115	1306	323	75
107	1116	1308	457	891
108	1117	1311	258	674
109	1118	1315	242	823
110	1119	1317	82	435
111	1120	1319	781	3306
112	1121	1323	1402	1671
113	1122	1329	279	665
114	1123	1336	37	765
115	1124	1337	177	389
116	1125	1338	887	744
117	1126	1339	248	724
118	1127	1341	298	525
119	1128	1342	26	445
120	1129	1344	23	370
121	1130	1345	160	402
122	1131	1351	2737	2600
123	1132	1353	655	792
124	1133	1354	94	354
125	1134	1356	679	849
126	1135	1358	679	849
127	1136	1359	32	346
128	1137	1361	271	426
129	1138	1362	637	1197
130	1139	1363	24	350
131	1140	1364	119	367
132	1141	1368	111	284
133	1142	1377	1221	1358
134	1143	1378	643	470
135	1144	138	99	539
136	1145	1382	994	686
137	1146	1384	34	264
138	1147	1386	124	477
139	1148	1389	1197	1
140	1149	139	94	294
141	1150	1390	1262	1053
142	1151	1393	1182	1325
143	1152	1394	1351	1542
144	1153	1395	229	411
145	1154	1396	923	1147
146	1155	1397	49	252
147	1156	1398	684	863
148	1157	1399	2613	286
149	1158	14	997	758
150	1159	1403	396	1

TABLE 3

SEQ ID NO: OF NUCLEOTIDE	SEQ ID NO: OF AMINO ACID	SEQ ID NO: IN USSN 09/491,404	START NUCLEOTIDE OF CODING REGION	STOP NUCLEOTIDE OF CODING REGION
151	1160	1406	735	1235
152	1161	1407	967	716
153	1162	1408	75	314
154	1163	1409	101	313
155	1164	141	384	551
156	1165	1414	242	532
157	1166	142	158	15
158	1167	1421	604	1425
159	1168	1422	1146	1835
160	1169	1423	2657	3295
161	1170	1424	315	163
162	1171	1426	39	509
163	1172	1427	892	686
164	1173	1428	395	619
165	1174	1430	284	514
166	1175	1432	178	2
167	1176	1433	1136	972
168	1177	1435	1283	1540
169	1178	1436	1669	2235
170	1179	144	55	219
171	1180	1440	363	121
172	1181	1441	1991	2197
173	1182	1443	1765	3054
174	1183	1445	1023	865
175	1184	1446	5692	5859
176	1185	1447	2959	2078
177	1186	1448	775	945
178	1187	1451	858	1430
179	1188	1453	1370	723
180	1189	1455	480	1007
181	1190	1457	278	451
182	1191	1459	824	561
183	1192	1460	56	463
184	1193	1461	184	480
185	1194	1462	486	635
186	1195	1465	319	492
187	1196	1466	398	3
188	1197	1468	262	453
189	1198	1476	526	684
190	1199	148	271	420
191	1200	1482	568	714
192	1201	1484	203	340
193	1202	1486	2185	1190
194	1203	1492	438	2912
195	1204	1493	82	225
196	1205	1501	210	347
197	1206	1508	1364	1101
198	1207	1509	56	613
199	1208	1512	828	965
200	1209	1515	3216	3812

TABLE 3

SEQ ID NO: OF NUCLEOTIDE	SEQ ID NO: OF AMINO ACID	SEQ ID NO: IN USSN 09/491,404	START NUCLEOTIDE OF CODING REGION	STOP NUCLEOTIDE OF CODING REGION
201	1210	1516	614	790
202	1211	1522	1709	1029
203	1212	1524	614	799
204	1213	1526	3917	4081
205	1214	1529	221	2146
206	1215	1530	644	390
207	1216	1532	16	1224
208	1217	1535	885	1031
209	1218	1536	245	1156
210	1219	1538	1617	4994
211	1220	154	97	234
212	1221	1540	4325	4158
213	1222	1541	2020	2778
214	1223	1544	595	3168
215	1224	1545	328	534
216	1225	1548	47	211
217	1226	1550	49	201
218	1227	1552	418	558
219	1228	1555	509	330
220	1229	1557	699	854
221	1230	1561	847	1932
222	1231	1563	775	933
223	1232	1565	286	453
224	1233	1567	807	974
225	1234	1568	1227	1601
226	1235	1569	113	328
227	1236	157	145	2
228	1237	1570	222	845
229	1238	1572	167	685
230	1239	1574	97	1167
231	1240	1575	581	2701
232	1241	1577	1246	953
233	1242	1578	1440	175
234	1243	1579	4738	4601
235	1244	1580	1431	1568
236	1245	1581	2491	3222
237	1246	1584	463	2157
238	1247	1585	156	2366
239	1248	1586	167	691
240	1249	1587	102	305
241	1250	1589	1157	1783
242	1251	159	812	639
243	1252	1592	270	521
244	1253	1593	92	310
245	1254	1594	814	188
246	1255	1595	101	2290
247	1256	1597	119	910
248	1257	1598	178	1398
249	1258	1600	2937	2578
250	1259	1604	47	526

TABLE 3

SEQ ID NO: OF NUCLEOTIDE	SEQ ID NO: OF AMINO ACID	SEQ ID NO: IN USSN 09/491,404	START NUCLEOTIDE OF CODING REGION	STOP NUCLEOTIDE OF CODING REGION
251	1260	1606	2204	1872
252	1261	1608	235	603
253	1262	1609	156	2366
254	1263	1611	1992	2135
255	1264	1614	968	786
256	1265	1615	2578	2751
257	1266	1616	6256	5813
258	1267	1617	29	709
259	1268	1619	1123	4071
260	1269	1621	581	2704
261	1270	1626	43	321
262	1271	1629	3616	1673
263	1272	163	509	183
264	1273	1630	81	248
265	1274	1631	9	572
266	1275	1633	2565	2807
267	1276	1634	2373	2510
268	1277	1635	3216	4508
269	1278	1636	4239	4081
270	1279	1642	4238	4020
271	1280	1643	152	304
272	1281	1644	47	478
273	1282	1645	121	921
274	1283	1646	3815	3030
275	1284	1647	335	186
276	1285	1649	6	974
277	1286	1654	34	951
278	1287	1655	491	1387
279	1288	1656	78	560
280	1289	1657	1431	1568
281	1290	1658	2373	1015
282	1291	1670	236	3
283	1292	1673	95	1342
284	1293	1685	2124	1786
285	1294	1690	245	415
286	1295	1691	977	774
287	1296	1699	50	247
288	1297	17	282	112
289	1298	1710	943	239
290	1299	1711	127	318
291	1300	1718	99	338
292	1301	1719	122	382
293	1302	172	33	461
294	1303	1720	180	1
295	1304	1722	160	327
296	1305	1726	175	363
297	1306	1737	84	497
298	1307	1738	188	379
299	1308	174	138	332
300	1309	1743	560	784

TABLE 3

SEQ ID NO: OF NUCLEOTIDE	SEQ ID NO: OF AMINO ACID	SEQ ID NO: IN USSN 09/491,404	START NUCLEOTIDE OF CODING REGION	STOP NUCLEOTIDE OF CODING REGION
301	1310	1747	1824	1961
302	1311	1748	97	411
303	1312	1749	151	492
304	1313	177	59	322
305	1314	1776	68	262
306	1315	1779	43	255
307	1316	178	58	399
308	1317	1781	1179	907
309	1318	1786	579	385
310	1319	1789	56	193
311	1320	180	218	78
312	1321	1800	230	394
313	1322	1801	1778	876
314	1323	181	174	428
315	1324	1829	179	42
316	1325	1846	525	785
317	1326	1848	5632	5838
318	1327	185	92	400
319	1328	1850	178	333
320	1329	186	699	1310
321	1330	1860	8	604
322	1331	1868	376	618
323	1332	187	148	366
324	1333	1870	233	388
325	1334	1872	12	206
326	1335	188	181	516
327	1336	1884	549	863
328	1337	1886	128	298
329	1338	189	28	204
330	1339	1891	11246	11097
331	1340	1895	175	417
332	1341	1897	221	400
333	1342	1899	744	890
334	1343	191	77	286
335	1344	1914	403	699
336	1345	192	8	343
337	1346	1947	656	1735
338	1347	1948	32	283
339	1348	195	129	323
340	1349	196	122	295
341	1350	1962	554	733
342	1351	197	110	277
343	1352	1976	348	2450
344	1353	198	93	239
345	1354	1980	137	310
346	1355	2	916	13698
347	1356	20	112	303
348	1357	2005	88	420
349	1358	2007	525	385
350	1359	2008	266	484

TABLE 3

SEQ ID NO: OF NUCLEOTIDE	SEQ ID NO: OF AMINO ACID	SEQ ID NO: IN USSN 09/491,404	START NUCLEOTIDE OF CODING REGION	STOP NUCLEOTIDE OF CODING REGION
351	1360	2013	64	234
352	1361	2016	99	329
353	1362	2018	84	401
354	1363	202	300	130
355	1364	2022	1240	1016
356	1365	2029	191	364
357	1366	2037	231	404
358	1367	2043	3206	3349
359	1368	2047	169	456
360	1369	2048	295	522
361	1370	2049	533	769
362	1371	205	4	684
363	1372	2051	403	699
364	1373	2055	173	379
365	1374	2056	270	1157
366	1375	2061	949	725
367	1376	2064	127	309
368	1377	2065	248	577
369	1378	2070	204	344
370	1379	2071	374	793
371	1380	2074	945	796
372	1381	2076	300	67
373	1382	2078	416	586
374	1383	2081	316	507
375	1384	2082	20	220
376	1385	209	19	168
377	1386	210	27	395
378	1387	2102	258	452
379	1388	2104	1706	1539
380	1389	211	84	311
381	1390	212	677	231
382	1391	2120	40	414
383	1392	214	101	268
384	1393	2140	213	377
385	1394	2161	216	368
386	1395	2162	106	420
387	1396	2164	104	250
388	1397	217	333	22
389	1398	218	80	325
390	1399	219	709	506
391	1400	2196	158	319
392	1401	2198	469	1164
393	1402	22	843	700
394	1403	2214	980	822
395	1404	2215	49	318
396	1405	2225	544	1974
397	1406	223	185	21
398	1407	2233	116	313
399	1408	224	189	16
400	1409	2240	2740	2525

TABLE 3

SEQ ID NO: OF NUCLEOTIDE	SEQ ID NO: OF AMINO ACID	SEQ ID NO: IN USSN 09/491,404	START NUCLEOTIDE OF CODING REGION	STOP NUCLEOTIDE OF CODING REGION
401	1410	2244	1489	1647
402	1411	2254	72	317
403	1412	226	335	120
404	1413	2260	562	738
405	1414	2268	300	67
406	1415	227	103	615
407	1416	2273	114	344
408	1417	2275	239	985
409	1418	2276	1358	1164
410	1419	2288	56	1459
411	1420	2291	83	532
412	1421	2296	264	530
413	1422	2298	533	781
414	1423	2300	1684	1845
415	1424	2305	8	226
416	1425	231	86	820
417	1426	232	361	1920
418	1427	233	150	467
419	1428	2331	334	2856
420	1429	2334	168	953
421	1430	2341	198	395
422	1431	2344	122	1432
423	1432	2346	1345	1187
424	1433	2348	502	729
425	1434	235	338	844
426	1435	2351	228	713
427	1436	236	232	2
428	1437	2360	1611	1357
429	1438	2362	36	263
430	1439	2364	294	1568
431	1440	2365	103	312
432	1441	2378	209	5281
433	1442	238	53	511
434	1443	2380	207	380
435	1444	239	457	663
436	1445	2392	176	2653
437	1446	2399	940	2040
438	1447	2405	144	380
439	1448	2407	1875	2702
440	1449	2415	1927	137
441	1450	242	1813	986
442	1451	2421	43	405
443	1452	2423	1556	1413
444	1453	2424	673	1041
445	1454	2432	295	1275
446	1455	2438	607	437
447	1456	2444	294	437
448	1457	2447	212	1588
449	1458	2448	52	1440
450	1459	2449	637	1197

TABLE 3

SEQ ID NO: OF NUCLEOTIDE	SEQ ID NO: OF AMINO ACID	SEQ ID NO: IN USSN 09/491,404	START NUCLEOTIDE OF CODING REGION	STOP NUCLEOTIDE OF CODING REGION
451	1460	245	208	876
452	1461	2450	3740	4369
453	1462	2453	222	389
454	1463	246	566	763
455	1464	2466	179	778
456	1465	2471	532	669
457	1466	2473	817	650
458	1467	2474	236	1333
459	1468	2476	173	3
460	1469	248	331	2
461	1470	2486	709	885
462	1471	249	88	456
463	1472	2496	107	1054
464	1473	2498	413	607
465	1474	2501	103	267
466	1475	2503	334	717
467	1476	2506	3740	4369
468	1477	2509	188	18
469	1478	2512	78	368
470	1479	2514	16	354
471	1480	2523	53	325
472	1481	2526	223	384
473	1482	2532	596	763
474	1483	2533	62	667
475	1484	2535	89	1519
476	1485	2537	175	375
477	1486	254	299	21
478	1487	2540	553	816
479	1488	2546	1905	1102
480	1489	2555	2046	4541
481	1490	2559	569	733
482	1491	256	9	410
483	1492	2560	288	76
484	1493	2565	3269	3502
485	1494	2569	116	478
486	1495	257	203	475
487	1496	2571	2763	2548
488	1497	2572	65	652
489	1498	2575	70	294
490	1499	2576	1195	1010
491	1500	258	434	21
492	1501	2580	155	400
493	1502	2591	53	214
494	1503	2592	163	348
495	1504	26	261	398
496	1505	2605	277	420
497	1506	261	29	598
498	1507	2614	1331	1510
499	1508	2617	235	378
500	1509	262	204	458

TABLE 3

SEQ ID NO: OF NUCLEOTIDE	SEQ ID NO: OF AMINO ACID	SEQ ID NO: IN USSN 09/491,404	START NUCLEOTIDE OF CODING REGION	STOP NUCLEOTIDE OF CODING REGION
501	1510	2624	254	418
502	1511	263	247	570
503	1512	264	184	540
504	1513	2643	1108	4026
505	1514	2644	305	535
506	1515	2645	1952	1509
507	1516	2647	1225	404
508	1517	2648	41	778
509	1518	265	53	418
510	1519	2650	190	936
511	1520	2658	1576	2451
512	1521	2659	44	430
513	1522	266	350	153
514	1523	2663	785	1177
515	1524	2665	395	550
516	1525	2666	41	778
517	1526	2667	244	384
518	1527	2668	174	527
519	1528	2669	27	302
520	1529	2678	1172	960
521	1530	2684	178	432
522	1531	269	341	520
523	1532	2699	1241	1083
524	1533	2701	402	2624
525	1534	2702	28	177
526	1535	2706	1108	4026
527	1536	2707	1240	1016
528	1537	271	59	346
529	1538	2714	34	987
530	1539	2715	1117	647
531	1540	2717	25	429
532	1541	2718	1670	1885
533	1542	2719	31	1137
534	1543	272	6	152
535	1544	2726	230	592
536	1545	2728	578	369
537	1546	2731	193	366
538	1547	2735	495	301
539	1548	274	352	119
540	1549	2741	94	255
541	1550	2798	1031	1240
542	1551	28	54	725
543	1552	2803	204	374
544	1553	2809	216	938
545	1554	2822	280	447
546	1555	2823	197	388
547	1556	2824	224	12
548	1557	2826	79	456
549	1558	2828	24	428
550	1559	2838	90	698

TABLE 3

SEQ ID NO: OF NUCLEOTIDE	SEQ ID NO: OF AMINO ACID	SEQ ID NO: IN USSN 09/491,404	START NUCLEOTIDE OF CODING REGION	STOP NUCLEOTIDE OF CODING REGION
551	1560	284	21	197
552	1561	2847	113	262
553	1562	285	146	292
554	1563	2852	233	439
555	1564	2854	830	988
556	1565	2855	336	1043
557	1566	2856	384	614
558	1567	2857	437	748
559	1568	2859	1295	1158
560	1569	286	30	179
561	1570	2860	2618	2469
562	1571	2864	1325	1176
563	1572	2867	1034	795
564	1573	288	190	345
565	1574	2884	856	257
566	1575	2886	15	167
567	1576	2891	34	405
568	1577	2900	104	2683
569	1578	2901	193	366
570	1579	2902	91	1806
571	1580	2907	268	498
572	1581	2908	83	1564
573	1582	2910	2131	3117
574	1583	2915	715	861
575	1584	2916	52	2064
576	1585	2919	62	1015
577	1586	292	615	854
578	1587	2923	332	1279
579	1588	2924	264	422
580	1589	2925	122	1432
581	1590	2930	195	341
582	1591	2931	221	3
583	1592	2934	1642	1827
584	1593	2937	38	421
585	1594	2940	520	383
586	1595	2944	325	68
587	1596	295	49	255
588	1597	2950	226	59
589	1598	2951	110	400
590	1599	2955	303	641
591	1600	2957	365	673
592	1601	2964	96	347
593	1602	2967	738	466
594	1603	2968	222	428
595	1604	2969	365	117
596	1605	2970	314	643
597	1606	2973	961	1176
598	1607	2975	975	799
599	1608	2979	89	442
600	1609	298	152	3

TABLE 3

SEQ ID NO: OF NUCLEOTIDE	SEQ ID NO: OF AMINO ACID	SEQ ID NO: IN USSN 09/491,404	START NUCLEOTIDE OF CODING REGION	STOP NUCLEOTIDE OF CODING REGION
601	1610	2991	112	261
602	1611	2995	201	368
603	1612	3	13559	13335
604	1613	30	176	751
605	1614	3002	1807	2265
606	1615	3005	339	743
607	1616	3023	64	243
608	1617	3039	71	217
609	1618	304	50	334
610	1619	305	226	387
611	1620	3051	56	268
612	1621	307	9	278
613	1622	308	116	274
614	1623	3085	97	3030
615	1624	3088	801	634
616	1625	3089	18	455
617	1626	3094	92	1246
618	1627	3098	40	342
619	1628	310	142	354
620	1629	3101	48	383
621	1630	3105	188	328
622	1631	3107	177	413
623	1632	3109	184	327
624	1633	3114	70	243
625	1634	3115	295	459
626	1635	3116	115	348
627	1636	3119	70	222
628	1637	3120	163	531
629	1638	3122	60	266
630	1639	3129	226	501
631	1640	3146	190	363
632	1641	3151	212	1588
633	1642	3153	86	517
634	1643	3165	244	453
635	1644	317	97	342
636	1645	3179	106	873
637	1646	3181	108	896
638	1647	3182	554	775
639	1648	3192	268	441
640	1649	3194	923	1192
641	1650	3195	38	376
642	1651	32	185	334
643	1652	3200	199	561
644	1653	3201	516	848
645	1654	3202	232	681
646	1655	3208	836	633
647	1656	3210	202	384
648	1657	3214	349	588
649	1658	3215	859	380
650	1659	3216	51	320

TABLE 3

SEQ ID NO: OF NUCLEOTIDE	SEQ ID NO: OF AMINO ACID	SEQ ID NO: IN USSN 09/491,404	START NUCLEOTIDE OF CODING REGION	STOP NUCLEOTIDE OF CODING REGION
651	1660	3220	116	283
652	1661	3222	324	545
653	1662	3227	385	1197
654	1663	323	65	223
655	1664	3240	385	1197
656	1665	3243	65	916
657	1666	3250	263	463
658	1667	3252	244	480
659	1668	3253	136	297
660	1669	3254	83	439
661	1670	3255	573	920
662	1671	3257	548	757
663	1672	3259	34	822
664	1673	326	58	525
665	1674	3263	102	350
666	1675	3270	313	152
667	1676	3271	117	473
668	1677	3272	44	190
669	1678	3273	106	486
670	1679	3274	246	392
671	1680	3278	174	1
672	1681	3281	988	1134
673	1682	3282	101	334
674	1683	3291	129	284
675	1684	3294	101	595
676	1685	3296	107	565
677	1686	3298	130	552
678	1687	3299	333	515
679	1688	3300	324	121
680	1689	3303	378	157
681	1690	3306	296	637
682	1691	3307	1454	1660
683	1692	3309	163	471
684	1693	3311	335	478
685	1694	3312	5	280
686	1695	3313	298	546
687	1696	3314	50	526
688	1697	3315	99	413
689	1698	3322	101	685
690	1699	3323	66	356
691	1700	3324	76	462
692	1701	3328	248	904
693	1702	3335	136	393
694	1703	3336	47	733
695	1704	3338	181	786
696	1705	3339	58	231
697	1706	3342	226	390
698	1707	3349	72	488
699	1708	3356	208	384
700	1709	3358	194	436

TABLE 3

SEQ ID NO: OF NUCLEOTIDE	SEQ ID NO: OF AMINO ACID	SEQ ID NO: IN USSN 09/491,404	START NUCLEOTIDE OF CODING REGION	STOP NUCLEOTIDE OF CODING REGION
701	1710	3360	263	1459
702	1711	3366	55	816
703	1712	3367	364	735
704	1713	3370	237	878
705	1714	3371	188	721
706	1715	3372	14	241
707	1716	3373	42	290
708	1717	3387	32	202
709	1718	3389	29	256
710	1719	3390	181	393
711	1720	3396	520	822
712	1721	3410	10	153
713	1722	3412	82	291
714	1723	3414	453	292
715	1724	3421	158	337
716	1725	3427	430	618
717	1726	3430	210	380
718	1727	3431	295	432
719	1728	3440	419	556
720	1729	3444	402	256
721	1730	3445	281	430
722	1731	346	42	722
723	1732	347	384	689
724	1733	3470	114	530
725	1734	3478	38	217
726	1735	3479	161	379
727	1736	348	37	231
728	1737	3482	156	296
729	1738	35	255	575
730	1739	3503	185	454
731	1740	3505	252	422
732	1741	3529	37	183
733	1742	353	262	522
734	1743	3537	127	273
735	1744	3539	98	268
736	1745	3542	25	312
737	1746	3543	70	228
738	1747	3544	31	177
739	1748	3548	972	385
740	1749	3553	27	164
741	1750	3560	113	358
742	1751	3563	483	764
743	1752	3564	6	434
744	1753	3566	316	507
745	1754	3570	6	377
746	1755	3574	108	440
747	1756	3576	569	348
748	1757	3579	293	442
749	1758	3582	20	388
750	1759	3583	172	396

TABLE 3

SEQ ID NO: OF NUCLEOTIDE	SEQ ID NO: OF AMINO ACID	SEQ ID NO: IN USSN 09/491,404	START NUCLEOTIDE OF CODING REGION	STOP NUCLEOTIDE OF CODING REGION
751	1760	3587	84	449
752	1761	3596	91	459
753	1762	3599	40	474
754	1763	3606	335	1105
755	1764	3609	169	666
756	1765	3617	141	410
757	1766	3620	218	388
758	1767	3630	189	1
759	1768	3642	122	643
760	1769	3644	431	664
761	1770	3647	274	720
762	1771	3651	245	472
763	1772	3652	259	642
764	1773	3653	153	1994
765	1774	3654	87	554
766	1775	3657	57	2744
767	1776	3658	387	920
768	1777	366	402	578
769	1778	3660	120	530
770	1779	3661	480	674
771	1780	3663	1096	938
772	1781	3669	689	1015
773	1782	3677	469	642
774	1783	3678	1194	889
775	1784	3685	406	1134
776	1785	3689	233	706
777	1786	3693	21	446
778	1787	3699	55	414
779	1788	370	59	262
780	1789	3707	38	436
781	1790	3711	229	474
782	1791	3713	314	463
783	1792	3717	178	675
784	1793	3720	258	695
785	1794	3721	96	548
786	1795	3722	32	562
787	1796	3724	220	513
788	1797	3726	180	467
789	1798	3729	251	523
790	1799	373	110	340
791	1800	3735	91	636
792	1801	3736	275	880
793	1802	3738	106	621
794	1803	3762	702	1175
795	1804	3768	293	598
796	1805	377	96	257
797	1806	3772	169	2
798	1807	3786	108	248
799	1808	3787	282	638
800	1809	3789	139	411

TABLE 3

SEQ ID NO: OF NUCLEOTIDE	SEQ ID NO: OF AMINO ACID	SEQ ID NO: IN USSN 09/491,404	START NUCLEOTIDE OF CODING REGION	STOP NUCLEOTIDE OF CODING REGION
801	1810	379	248	421
802	1811	38	146	3
803	1812	382	24	275
804	1813	385	138	1
805	1814	388	268	74
806	1815	39	302	3
807	1816	391	24	368
808	1817	395	51	482
809	1818	397	422	766
810	1819	399	102	311
811	1820	4	11219	13123
812	1821	405	253	2
813	1822	406	342	665
814	1823	411	321	542
815	1824	416	736	909
816	1825	422	1541	867
817	1826	43	330	686
818	1827	434	207	34
819	1828	435	140	445
820	1829	437	160	423
821	1830	439	347	706
822	1831	44	91	282
823	1832	450	136	402
824	1833	458	169	348
825	1834	459	99	284
826	1835	462	70	282
827	1836	465	462	791
828	1837	467	76	348
829	1838	470	35	637
830	1839	475	37	426
831	1840	477	242	382
832	1841	478	66	311
833	1842	485	196	426
834	1843	488	117	443
835	1844	490	231	485
836	1845	493	281	610
837	1846	496	90	371
838	1847	5	34	3933
839	1848	501	60	368
840	1849	502	707	856
841	1850	504	208	459
842	1851	505	165	317
843	1852	509	62	223
844	1853	511	46	432
845	1854	515	13	582
846	1855	516	92	325
847	1856	518	83	283
848	1857	519	365	685
849	1858	521	12	413
850	1859	525	6	251

TABLE 3

SEQ ID NO: OF NUCLEOTIDE	SEQ ID NO: OF AMINO ACID	SEQ ID NO: IN USSN 09/491,404	START NUCLEOTIDE OF CODING REGION	STOP NUCLEOTIDE OF CODING REGION
851	1860	526	862	725
852	1861	532	207	590
853	1862	536	226	53
854	1863	537	49	198
855	1864	540	270	1
856	1865	541	38	412
857	1866	546	388	2
858	1867	555	199	438
859	1868	556	144	482
860	1869	559	380	165
861	1870	563	27	617
862	1871	566	158	382
863	1872	568	69	320
864	1873	57	6	158
865	1874	571	8	1516
866	1875	572	32	505
867	1876	573	139	456
868	1877	574	49	771
869	1878	576	519	370
870	1879	578	168	1
871	1880	580	159	641
872	1881	581	108	497
873	1882	582	80	403
874	1883	587	172	435
875	1884	589	27	374
876	1885	590	84	428
877	1886	595	68	1138
878	1887	598	1023	766
879	1888	61	65	208
880	1889	612	310	546
881	1890	614	166	918
882	1891	617	252	602
883	1892	62	969	661
884	1893	620	188	418
885	1894	622	877	1014
886	1895	629	202	687
887	1896	63	98	277
888	1897	632	221	367
889	1898	64	536	381
890	1899	640	338	3
891	1900	641	12	395
892	1901	642	194	397
893	1902	644	15	395
894	1903	646	132	380
895	1904	647	3	389
896	1905	650	135	413
897	1906	651	231	428
898	1907	653	128	442
899	1908	654	214	77
900	1909	656	49	465

TABLE 3

SEQ ID NO: OF NUCLEOTIDE	SEQ ID NO: OF AMINO ACID	SEQ ID NO: IN USSN 09/491,404	START NUCLEOTIDE OF CODING REGION	STOP NUCLEOTIDE OF CODING REGION
901	1910	657	86	397
902	1911	66	267	614
903	1912	662	387	701
904	1913	666	76	498
905	1914	667	517	2184
906	1915	668	1423	788
907	1916	67	107	622
908	1917	678	172	387
909	1918	68	78	341
910	1919	680	832	671
911	1920	683	505	164
912	1921	687	105	521
913	1922	690	139	294
914	1923	691	244	456
915	1924	699	194	754
916	1925	701	371	520
917	1926	702	1888	2028
918	1927	704	1254	808
919	1928	705	126	1463
920	1929	706	31	390
921	1930	707	367	2
922	1931	709	1152	934
923	1932	715	744	541
924	1933	716	1360	1220
925	1934	722	173	430
926	1935	725	498	271
927	1936	727	18	164
928	1937	729	230	3
929	1938	73	262	834
930	1939	731	491	246
931	1940	740	20	322
932	1941	741	1430	1167
933	1942	747	660	523
934	1943	749	263	727
935	1944	750	209	391
936	1945	751	753	517
937	1946	755	172	387
938	1947	756	209	376
939	1948	76	656	513
940	1949	760	131	538
941	1950	763	893	1126
942	1951	766	1271	1537
943	1952	771	458	318
944	1953	775	391	558
945	1954	781	410	1684
946	1955	791	967	1284
947	1956	793	554	970
948	1957	795	8	268
949	1958	796	342	199
950	1959	798	211	405

TABLE 3

SEQ ID NO: OF NUCLEOTIDE	SEQ ID NO: OF AMINO ACID	SEQ ID NO: IN USSN 09/491,404	START NUCLEOTIDE OF CODING REGION	STOP NUCLEOTIDE OF CODING REGION
951	1960	799	625	392
952	1961	8	1523	1293
953	1962	801	484	678
954	1963	802	331	489
955	1964	808	210	905
956	1965	812	162	920
957	1966	819	723	2669
958	1967	820	964	725
959	1968	825	182	328
960	1969	829	1843	2292
961	1970	830	58	201
962	1971	832	150	341
963	1972	835	130	762
964	1973	836	449	291
965	1974	838	175	324
966	1975	84	175	435
967	1976	842	73	393
968	1977	844	423	824
969	1978	845	214	32
970	1979	846	120	317
971	1980	847	212	364
972	1981	85	190	426
973	1982	852	74	541
974	1983	855	1653	1465
975	1984	857	1964	2659
976	1985	858	598	1020
977	1986	861	58	933
978	1987	876	222	779
979	1988	878	2021	2161
980	1989	879	189	362
981	1990	88	39	278
982	1991	886	1165	1022
983	1992	891	158	310
984	1993	892	759	995
985	1994	895	224	379
986	1995	897	131	622
987	1996	9	1678	1448
988	1997	901	55	753
989	1998	906	450	623
990	1999	913	40	237
991	2000	918	17	334
992	2001	92	385	122
993	2002	926	772	518
994	2003	929	146	283
995	2004	932	23	175
996	2005	934	38	235
997	2006	935	286	423
998	2007	936	24	284
999	2008	939	450	623
1000	2009	94	139	2

TABLE 3

SEQ ID NO: OF NUCLEOTIDE	SEQ ID NO: OF AMINO ACID	SEQ ID NO: IN USSN 09/491,404	START NUCLEOTIDE OF CODING REGION	STOP NUCLEOTIDE OF CODING REGION
1001	2010	944	156	860
1002	2011	947	174	356
1003	2012	957	80	400
1004	2013	96	187	387
1005	2014	964	1352	1528
1006	2015	97	166	2
1007	2016	98	535	344
1008	2017	995	559	386
1009	2018	997	34	231

## WHAT IS CLAIMED IS:

1. An isolated polynucleotide comprising a nucleotide sequence selected from the group consisting of SEQ ID NO: 1-1009, a mature protein coding portion of SEQ ID NO: 1-1009, an active domain of SEQ ID NO: 1-1009, and complementary sequences thereof.
2. An isolated polynucleotide encoding a polypeptide with biological activity, wherein said polynucleotide hybridizes to the polynucleotide of claim 1 under stringent hybridization conditions.
3. An isolated polynucleotide encoding a polypeptide with biological activity, wherein said polynucleotide has greater than about 90% sequence identity with the polynucleotide of claim 1.
4. The polynucleotide of claim 1 wherein said polynucleotide is DNA.
5. An isolated polynucleotide of claim 1 wherein said polynucleotide comprises the complementary sequences.
6. A vector comprising the polynucleotide of claim 1.
7. An expression vector comprising the polynucleotide of claim 1.
8. A host cell genetically engineered to comprise the polynucleotide of claim 1.
9. A host cell genetically engineered to comprise the polynucleotide of claim 1 operatively associated with a regulatory sequence that modulates expression of the polynucleotide in the host cell.
10. An isolated polypeptide, wherein the polypeptide is selected from the group consisting of:
  - (a) a polypeptide encoded by any one of the polynucleotides of claim 1; and

(b) a polypeptide encoded by a polynucleotide hybridizing under stringent conditions with any one of SEQ ID NO:1-1009.

11. A composition comprising the polypeptide of claim 10 and a carrier.

12. An antibody directed against the polypeptide of claim 10.

13. A method for detecting the polynucleotide of claim 1 in a sample, comprising:

- contacting the sample with a compound that binds to and forms a complex with the polynucleotide of claim 1 for a period sufficient to form the complex; and
- detecting the complex, so that if a complex is detected, the polynucleotide of claim 1 is detected.

14. A method for detecting the polynucleotide of claim 1 in a sample, comprising:

- contacting the sample under stringent hybridization conditions with nucleic acid primers that anneal to the polynucleotide of claim 1 under such conditions;
- amplifying a product comprising at least a portion of the polynucleotide of claim 1; and
- detecting said product and thereby the polynucleotide of claim 1 in the sample.

15. The method of claim 14, wherein the polynucleotide is an RNA molecule and the method further comprises reverse transcribing an annealed RNA molecule into a cDNA polynucleotide.

16. A method for detecting the polypeptide of claim 10 in a sample, comprising:

- contacting the sample with a compound that binds to and forms a complex with the polypeptide under conditions and for a period sufficient to form the complex; and

b) detecting formation of the complex, so that if a complex formation is detected, the polypeptide of claim 10 is detected.

17. A method for identifying a compound that binds to the polypeptide of claim 10, comprising:

- a) contacting the compound with the polypeptide of claim 10 under conditions sufficient to form a polypeptide/compound complex; and
- b) detecting the complex, so that if the polypeptide/compound complex is detected, a compound that binds to the polypeptide of claim 10 is identified.

18. A method for identifying a compound that binds to the polypeptide of claim 10, comprising:

- a) contacting the compound with the polypeptide of claim 10, in a cell, under conditions sufficient to form a polypeptide/compound complex, wherein the complex drives expression of a reporter gene sequence in the cell; and
- b) detecting the complex by detecting reporter gene sequence expression, so that if the polypeptide/compound complex is detected, a compound that binds to the polypeptide of claim 10 is identified.

19. A method of producing the polypeptide of claim 10, comprising,

- a) culturing a host cell comprising a polynucleotide sequence selected from the group consisting of a polynucleotide sequence of SEQ ID NO: 1-1009, a mature protein coding portion of SEQ ID NO: 1-1009, an active domain of SEQ ID NO: 1-1009, complementary sequences thereof and a polynucleotide sequence hybridizing under stringent conditions to SEQ ID NO: 1-1009, under conditions sufficient to express the polypeptide in said cell; and
- b) isolating the polypeptide from the cell culture or cells of step (a).

20. An isolated polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO: 1010-2018, the mature protein portion thereof, or the active domain thereof.

21. The polypeptide of claim 20 wherein the polypeptide is provided on a polypeptide array.
22. A collection of polynucleotides, wherein the collection comprises the sequence information of at least one of SEQ ID NO: 1-1009.
23. The collection of claim 22, wherein the collection is provided on a nucleic acid array.
24. The collection of claim 23, wherein the array detects full-matches to any one of the polynucleotides in the collection.
25. The collection of claim 23, wherein the array detects mismatches to any one of the polynucleotides in the collection.
26. The collection of claim 22, wherein the collection is provided in a computer-readable format.
27. A method of treatment comprising administering to a mammalian subject in need thereof a therapeutic amount of a composition comprising a polypeptide of claim 10 or 20 and a pharmaceutically acceptable carrier.
28. A method of treatment comprising administering to a mammalian subject in need thereof a therapeutic amount of a composition comprising an antibody that specifically binds to a polypeptide of claim 10 or 20 and a pharmaceutically acceptable carrier.